

**ACCURACY IN EMPATHIC AND TRAIT JUDGEMENTS:
THE QUEST FOR THE GOOD JUDGE, GOOD TARGET, GOOD
TRAIT, AND GOOD RELATIONSHIP**

**by
Geoff Thomas**

Submitted in fulfilment of the
requirements for the degree of
Doctor of Philosophy

Department of Psychology
University Of Canterbury
Christchurch, New Zealand
1999

F
98.4
F456
1999

DEDICATION

For my wife, Kirsty – you were my inspiration and support.

ACKNOWLEDGEMENTS

My supervisor, Garth Fletcher, has not only contributed greatly to the success of this dissertation, he has been an exemplary role model. His influence has left me with an appreciation of the many qualities that make one a great social psychologist.

Of course, I am grateful to the dating couples who gave up their time and energy to participate in this research. Without their willingness to be videotaped while discussing intimate relationship problems, to be watched by friends and strangers alike, and to share their intimate thoughts and feelings, this research would not have been possible.

Finally, I wish to thank my family and friends for their encouragement over the course of this thesis.

TABLE OF CONTENTS

	Page
DEDICATION	ii
ACKNOWLEDGEMENTS	iii
TABLE OF CONTENTS	iv
LIST OF TABLES	vi
LIST OF FIGURES	viii
LIST OF APPENDICES	ix
ABSTRACT	1
GENERAL INTRODUCTION	2
The History of Interpersonal Accuracy Research	4
The Investigation of Empathic Accuracy	9
General Aims of this Thesis	11
STUDY 1: THE MODERATORS OF EMPATHIC ACCURACY	14
The Good Relationship	15
The Good Target	20
The Good Judge	23
The Characteristics of the Good Judge	25
The Characteristics of the Good Relationship	28
Overview of Study 1	32
Method	33
Results	45
Discussion	69
Summary of Findings	69
Explaining Relationship-level Effects	70
Explaining Target Effects	75
Explaining Judge Effects	76
STUDY2: THE MODERATORS OF TRAIT ACCURACY AND CONSENSUS	80
Good Information	81
The Good Trait	88
Trait × Information	89
Overview of Study 2	90

Results	91
Discussion	100
Summary of Findings	100
Acquaintanceship	100
Trait Observability	103
Acquaintanceship × Observability	104
STUDY 3: THE LINK BETWEEN EMPATHIC & TRAIT ACCURACY	105
Results	107
Discussion	109
GENERAL DISCUSSION	111
Summary of the Major Results	111
Individual Differences in Judgement Accuracy	112
The Judge-Target Relationship and Judgement Accuracy	117
Caveats and Conclusions	123
REFERENCES	125

LIST OF TABLES

	Page
Table 1: Means and Standard Deviations for Male and Female Dating Partners as Perceivers, and Correlations Between Dating Partners	46
Table 2: Zero-order Correlations Between Empathic Accuracy Scores for the Dating Partners as Targets and the Friends and Strangers as Perceivers	47
Table 3: Zero-order Correlations Between Empathic Accuracy Scores Across the Male and Female Married Targets, for the Four Groups of Perceivers: Male Partners, Female Partners, Friends, and Strangers	47
Table 4: Means and Standard Deviations for Friends and Strangers as Perceivers	49
Table 5: Zero-order Correlations Between Perceivers' Reliance on Diagnostic Behaviour and Empathic Accuracy Attained by Dating Partners, Friends, and Strangers	53
Table 6: Target: Zero-order Empathic Accuracy Correlations Across Different Perceiver Groups With the Male and Female Dating Partners as Targets	53
Table 7: Perceiver: Correlations Between Perceiver's Empathic Accuracy for Dating Partner(s) and Perceiver's Empathic Accuracy for Stranger Married Couples	54
Table 8: Zero-order Correlations of Predictor Variables With Male and Female Dating Partners' Empathic Accuracy for Both Their Respective Partners and the Stranger Married Couples as the Dependent Variables	56
Table 9: Zero-order Correlations of Predictor Variables With Friends' and Strangers' Empathic Accuracy for Both the Dating Couple and the Stranger Married Couples as the Dependent Variables	58
Table 10: Regression Coefficients Derived From Hierarchical Multiple Regression Analysis With Male Dating Partners' Empathic Accuracy Scores for his Partner as the Dependent Variable	60
Table 11: Regression Coefficients Derived From Hierarchical Multiple Regression Analysis With Female Dating Partners' Empathic Accuracy Scores for her Partner as the Dependent Variable	60

Table 12: Regression Coefficients Derived From Two Simultaneous Multiple Regression Analyses With Strangers' Empathic Accuracy Scores for the Dating Couple and the Stranger Married Couples as the Two Dependent Variables	64
Table 13: Descriptive Statistics for All Interjudge Agreement Indices, Derived From Profile Correlations	93
Table 14: Means of Dispositional Variables at Each Level of Acquaintanceship	95
Table 15: Correlations Between Big-Five Factors and Interjudge Agreement Indices, Calculated Across Participants	99
Table 16: Zero-order Correlations of Male and Female Dating Partners' Empathic Accuracy Scores With Selected Interjudge Agreement Indices	107
Table 17: Zero-order Correlations of Selected Interjudge Agreement Indices With Friends' and Strangers' Empathic Accuracy Scores for Male and Female Dating Partners	108

LIST OF FIGURES

	Page
Figure 1: A diagrammatic representation of the perceiver and target relationships in phase 1 of the study.	39
Figure 2: A diagrammatic representation of the perceiver and target relationships in session 2 of the study.	40
Figure 3: A diagrammatic representation of the rater similarity judgements derived from the thought/feeling statements.	41
Figure 4: Mean empathic accuracy scores for partners as targets across male and female perceivers who were strangers, friends, or partners.	50
Figure 5: A path analysis depicting the relations between female dating partners' attributional complexity and female dating partners' empathic accuracy, as mediated by female dating partners' problem-specific disclosure to the friend.	62
Figure 6: A path analysis depicting the relations between friends' closeness to the female dating partner and friends' empathic accuracy for the dating couple, as mediated by female dating partners' problem-specific disclosure to the friend.	63
Figure 7: Friend's empathic accuracy for the dating couple as a function of disclosure to the friend by the female dating partner and attributional complexity.	67
Figure 8: Friend's empathic accuracy for the dating couple as a function of disclosure to the friend by the female dating partner and verbal intelligence.	67

LIST OF APPENDICES

	Page
Appendix A: Big-Five Factor Loadings for 20 Bipolar Scales	137

ABSTRACT

This research examined the moderators of judgement accuracy in three studies. Study 1 tested three moderators of empathic accuracy assessed during close relationship interaction: (1) the ability of the judge, (2) the readability of the target, and (3) the level of acquaintanceship. The design involved multiple perceivers judging multiple targets. Fifty dating couples reviewed videotapes of their prior problem-solving discussions and described both their own on-line cognitions and those of their partners. Fifty friends of the couples and 50 strangers later inferred the on-line cognitions of the same dating couples while observing the videotapes of their interactions. In addition, all three groups observed another videotape of two married couples' problem-solving interactions (strangers to the entire sample) and attempted the same empathic accuracy task. As predicted, the results revealed evidence for the good judge and good relationship, but not the good target. Increased relationship closeness was associated with higher accuracy. Both attributional complexity and verbal intelligence predicted empathic performance (but at different levels of acquaintanceship) and women were consistently superior to men. Several mediational models were tested, the results of which showed that female partners' problem-specific disclosure to the friend mediated the link between (a) female partners' attributional complexity and their empathic accuracy, and (b) friends' closeness to the female partner and their empathic accuracy.

Study 2 examined the influence of trait observability and the level of acquaintanceship, along with their unique interaction, on accuracy and consensus in trait judgements. Personality judgements of 100 targets (the fifty dating couples from study 1) were provided by the self, partners, friends, and strangers. As expected, greater levels of both trait observability and acquaintanceship were associated with more accurate and consensual trait judgements. Moreover, acquaintanceship interacted with observability such that trait visibility was an important determinant of accuracy and consensus for strangers but not for friends and partners.

Study 3 investigated the generality of judges' performance across judgement domains. The results showed no relationship between judges' accuracy in describing the targets' on-line cognitions (in study 1) and their accuracy and consensus in judging the same targets' personality profile (in study 2).

The results are interpreted within a Realistic Accuracy Model (Funder, 1995) and a social cognitive framework that exploits the distinction between theory and data-driven judgements.

GENERAL INTRODUCTION

“This remarkable capacity we possess to understand something of the character of another person ... is a precondition of social life” (Asch, 1946, p. 285).

As Asch's comment illustrates, person perception is a fundamental and ubiquitous feature of daily existence. Such judgements dictate people's decisions about who to avoid, trust, marry, divorce, employ, sack, institutionalise, parole, and elect. Yet our interest in each other is often driven by intrinsic, as well as pragmatic motives (Funder, 1995). People frequently adopt the role of amateur psychologist simply because they are curious about others. Indeed, the plethora of sidewalk cafes, confessional television programmes, and tabloid gossip about the rich and famous, can really only be explained by peoples intrinsic fascination with interpreting and explaining one another's behaviour (Funder, 1999).

There are several categories of interpersonal judgements, two of which are pertinent to this thesis. First, people make judgements of other's stable and enduring dispositions. Personality traits are used to describe people (Joshua is intelligent and assertive), explain their behaviour (Nancy didn't go to the party because she is shy), and predict future events (that plan won't work because Gary is too impulsive). Such trait terms are an integral part of the lay person's social cognition. Pioneering research by Allport and Odbert (1936) uncovered a total of 17, 953 trait descriptions in the unabridged English dictionary. More recently, psychologists have used factor analysis to reveal that this vast number of traits boils down to an overarching set of five trait categories (termed the Big-Five). Furthermore, there is evidence that people spontaneously use traits when describing others and that traits are central to storing information about others in long-term memory (see Trope & Higgins, 1993).

The second domain of social judgement investigated in this thesis is empathic inference; that is, the attributions to one another of transient thoughts and feelings during social interaction. In virtually every social encounter, people make judgements about the thoughts, feelings, and intentions that drive other's behaviour. Such empathic judgements often occur rapidly and automatically. However, empathic inferences may also be conscious and tied to in-depth cognition, as people extensively search for clues to

determine whether someone's compliment is genuine or to uncover another's true feelings hidden behind his or her deadpan facial expression (Hodges & Wegner, 1997). In any case, such empathic cognition is regarded as a *sine qua non* of human relationships, providing a bridge between the inner psychological experiences of one person and those of another (Levenson & Ruef, 1992; Smither, 1977).

Given the importance and pervasiveness of both empathic and personality judgements, one natural question concerns the extent to which people's perceptions actually correspond with reality, and what features enhance or diminish such accuracy? The issue of judgement accuracy is the overarching theme of this thesis. Although it is one of the oldest topics in social and personality psychology, interpersonal accuracy has often bedevilled researchers. In the next section, I provide an overview of the controversial and chequered history of interpersonal accuracy research, concentrating on personality judgements and empathic judgements (for more detailed reviews see Cook, 1979; Funder, 1999; Kenny & Albright, 1987). The various arguments are reviewed, in part, to establish that studying the accuracy of the social perceiver is both an important and scientifically respectable task.

The History of Interpersonal Accuracy Research

Encouraged by the success of standardised intelligence testing, early accuracy researchers eagerly focused their attention on social intelligence (Kenny, 1994). A lively research tradition developed, dominated by the search for the prototypical "good judge" of personality and his or her associated personality characteristics (for reviews, see Bruner & Tagiuri, 1954; Davis & Kraus, 1997; Taft, 1955).

The Demise of Accuracy Research

In the mid 1950's, however, the wheels fell off the accuracy wagon (Davis & Kraus, 1997). Cronbach published a devastating methodological critique, which all but silenced research on the accuracy of personality judgements for the next three decades (Cronbach, 1955; Gage & Cronbach, 1955). He showed that measuring accuracy was not simply a matter of computing a difference score between the target's self-rating on a trait (the criterion) and a judge's rating of the target on that trait (the judgement). The crux of

the problem was that such difference scores were contaminated by a set of measurement artifacts, including “elevation”, “differential elevation”, “stereotype accuracy”, and “assumed similarity”, all of which rendered difference scores as essentially uninterpretable (see Kenny & Albright, 1987). Both elevation and differential elevation refer to the possible influence of shared response styles between judge and target. If both judge and target coincidentally share tendencies to use a rating scale in the same way (e.g., positive response bias), then associated accuracy scores could be artifactually elevated. Stereotype accuracy, another Cronbachian confound, can artifactually inflate accuracy scores when a judge provides ratings that characterise the personality of the prototypical person and, fortuitously, the target resembles the average person. Finally, assumed similarity raises the possibility that accuracy scores are contaminated to the extent that judges merely project their own personality traits onto the target, and the judge and target happen to share similar personality traits.

Cronbach’s critique effectively brought into disrepute virtually all of the research on accuracy to that date. Cronbach did propose statistical methods for eliminating these confounds. However, his solutions were difficult to implement in the pre-computer era, and expressed in cryptic mathematical notation that seemed to bamboozle most researchers (Kenny & Albright, 1987). Unwittingly, accuracy research became earmarked as an unresearchable topic in social and personality psychology (Cline, 1964).

As if this methodological damage was not serious enough, the status of interpersonal accuracy research suffered three further attacks over the next three decades. First, the status of personality judgements, the vanguard of accuracy research, became the topic of a bitter and protracted debate in psychology. The controversy emanated from Mischel’s (1968) influential argument that compared to the powerful influence that situations have on behaviour, cross-situational consistency and the associated effect of personality on behaviour is minor. His review of prior research seemed to show that behavioural consistency was typically below the .30 level, leading Mischel to throw out the challenge to find correlations above this supposedly trivial level (see Kenrick & Funder, 1988 for a detailed review). In its most radical form, the brand of situationism promoted by Mischel dismissed personality as a construct that merely existed in the eye of the beholder. The popularity of these claims within psychology created a major barrier

to accuracy research. Clearly if personality does not even exist, then both lay and scientific judgements of personality are, by implication, utterly inaccurate.

Second, the growth of the constructivist (or relativist) approach in social psychology (e.g., Gergen, 1989) lead to a different kind of assault on accuracy research. The essence of the relativist critique is that the notion of accuracy mistakenly implies an objective independent reality to be perceived. According to this perspective, there is no absolute truth or reality. Consequently, it is not possible to determine the accuracy of people's judgements – all perceptions are merely constructions of reality and are therefore equally valid (Kruglanski, 1989). Other researchers have taken a slightly different but related tack, and claimed that even if personality does exist it is impossible to find a perfect and objective indicator of such truth (e.g., Cook, 1979, Jones, 1985). In other words, the formidable problems associated with criterion measurement renders the scientific study of accuracy all but impossible.

In the wake of Cronbach's critique, social psychologists hastily shifted their attention away from the accuracy and content of social judgements to the underlying psychological process of such judgements (Funder & West, 1993). However, psychology's fascination with accuracy did not remain dormant for long. While studying cognitive processes underlying social judgements, researchers apparently demonstrated that humans characteristically erred by falling well short of what the experimenters regarded as rational and normative approaches to decision-making. Many of these researchers in the 1970's and 1980's concluded, albeit inappropriately, that the social perceiver was essentially inept.

As the catalogue of errors documented by psychologists grew, the rhetoric used to describe the validity of human judgement became increasingly pessimistic. Descriptions such as "the hapless judge", "dramatic failures of the human mind", "pathologies of preference as well as of inference behaviour", and "the fallibility of judgement" became commonplace (Christensen-Szalanski & Beach, 1984; Fletcher, 1995; Hastie & Rasinski, 1988). Such failings were claimed to be particularly evident in the domain of lay personality judgement, which was viewed by many as so error-prone as to be essentially useless for any purpose outside of the study of error itself (Funder, 1983). Taken together all these conclusions conjure up an image of the social perceiver as so flawed that any reference to the accuracy of social judgements almost seems inane.

The Revival of Accuracy Research

The lack of research directly dealing with the accuracy of personality judgement was finally broken in the early 1980's when a few key psychologists mounted a defence of judgement accuracy. This first phase of the accuracy renaissance involved a vigorous attempt to demonstrate three key propositions: a) the existence and consistency of personality, b) that personality judgements are, at least in some cases, accurate according to reasonable criteria, and c) that given due care measures of judgement accuracy are not hopelessly contaminated by artifacts. I briefly outline each of these critical arguments.

During the last two decades, personality researchers have accumulated evidence to rebuff the challenges of the situationist and constructivist critiques. First, several studies have found substantial correlations (higher than the .30 ceiling) between questionnaire measures of personality and independent observations of behaviour (e.g., Block, Buss, Block & Gjerde, 1981; Moskowitz & Schwarz 1982), particularly when aggregate measures of behaviour were used to increase measurement reliability (Epstein & O'Brien, 1985). In fact, personality variables often generate similar effect sizes to situational variables when predicting behaviour (Funder & Ozer, 1983). Second, Funder and Colvin (1991) demonstrated that despite the presence of powerful situational forces, people still managed to substantially maintain their unique behavioural styles across situations. Hence, individual differences in personality do exist. Of course, the extent to which people can judge personality traits accurately is another matter.

Recently, there has been a growing recognition in social and personality psychology that the pendulum had swung too far towards an image of the inept and fundamentally flawed social perceiver (e.g., Fletcher, 1995; Lopes, 1991). The portrayal of the social perceiver as hopelessly inept is problematic for several reasons. First, from a pragmatic and evolutionary perspective, people must generally navigate their social world with at least a modicum of success. Not everyone is left by their spouse, raped by an acquaintance, or swindled by their accountant. Second, many researchers studying error have employed a null hypothesis of perfect accuracy, and when errors have occurred, have inappropriately concluded that people's judgements have no validity (Hastie & Rasinski, 1988). Third, the normative criteria typically used to assess how valid or rational lay social cognition is, have been increasingly called into question (see, for example, Fletcher, 1995).

Finally, there is good evidence that people's personality judgements often possess reasonable accuracy according to realistic criteria. Numerous studies have demonstrated significant agreement between people judging the personality traits of the same target (see Kenny, 1994). Such agreement occurs whether the comparison is between other's judgements and the target's self-judgement (hereafter referred to as *accuracy* or *self-other agreement*), or between different judges of the same target (hereafter referred to as *consensus*). Furthermore, as already mentioned, personality judgements have been found to predict, with considerable validity, direct observations of behaviour. This evidence for the existence of accurate personality judgement, however, begs the question of whether such realistic criteria comprise a sufficiently good measure of the target's true personality.

The criterion problem is by no means unique to accuracy research. Ever since Heisenberg, scientists have invariably accepted that there is an element of uncertainty and subjectivity in measurement. Therefore, sceptics who have spurned accuracy criterion for its fallibility have demanded a standard of proof that is obviously unreasonable (e.g., Jones, 1985). One way of solving this problem is to adopt a multifaceted approach to the evaluation of accuracy criterion. For example, according to Funder's (1995) Realistic Accuracy Model (RAM), the accuracy of personality judgements, like the validity of a questionnaire, cannot be established in relation to any single criterion, but only through the convergence of multiple sources of pertinent evidence concerning the target. For instance, self-other agreement, the most frequently used measure, seems a reasonable criterion for accuracy. If two judges disagree substantially, then at least one judge must be wrong. But it is not a sufficient criterion because two judges can agree but still be wrong. However, as a number of recent studies have shown, if self-other agreement is bolstered by other relevant criterion, such as consensus and behavioural prediction, that converge in meaningful ways (e.g., Colvin & Funder, 1991; Colvin, 1993b), then a persuasive case can be mounted for the valid measurement of accuracy of personality judgement.

The final major task for accuracy researchers was to deal with the serious methodological concerns raised by Cronbach (1955). Recall that Cronbach criticised the use of self-other agreement scores as being potentially so contaminated by measurement artifacts that they were uninterpretable. With the benefit of hindsight, these problems

were not intractable as was once thought. Indeed, with the development of computers, powerful statistical programmes, and associated statistical techniques, Cronbachian artifacts are no longer difficult to circumvent. The two elevation components can be dealt with either by the use of correlational measures or forced choice ratings techniques. The stereotype accuracy and assumed similarity components can be obviated by using appropriate statistical controls, or via research design, or both.

The Investigation of Empathic Accuracy

Along with some of the more general, previously described, concerns that apply to any form of judgement accuracy, the study of empathic accuracy has encountered its own vexing set of problems. One of the greatest sources of confusion has been the absence of a concise and consensual definition of empathic accuracy. There are three important distinctions that need to be made in order to differentiate empathic accuracy from related concepts. First, early researchers often referred to empathic and trait judgement accuracy as if they were synonymous, which they clearly are not. Second, the distinction between empathy (the process) and empathic accuracy (the outcome) has often been blurred. Empathy has been frequently referred to as a process of cognitive and affective perspective taking, whereas empathic accuracy signifies the success of this endeavour (see Davis, 1994, for a more detailed review). Finally, empathic accuracy has sometimes been confused with definitions that assert the sharing of affect between empathiser and target is the primary component of empathy. Empathic accuracy involves an understanding of the target's episodic psychological state, without necessarily registering a vicarious emotional response. For example, an attribution that "Louise is angry" does not necessarily involve the experience of anger in the attributer.

The measurement of empathic accuracy has also presented problems. Early attempts to assess empathic accuracy often relied on either some form of self-report scale designed to measure dispositional empathic ability (e.g., Davis, 1983, Hogan, 1969). Apart from possible bias arising from social desirability, the major obstacle to the use of self-report measures is that they assume that people possess accurate meta-knowledge about their own levels of empathic ability - an assumption that recent research has all but dismissed (Ickes, 1993). A second approach measured the perceiver's accuracy or

affective sensitivity in inferring the emotional tone of target's nonverbal cues (see Ickes, 1993; Taft, 1955, for a review). In the typical affective sensitivity study, targets feigned particular emotions while either being photographed or reading ambiguous statements, and the perceiver was required to choose the correct emotion from a set of prefabricated response options (e.g., Noller, 1980). The difficulty with such procedures concerns the extent to which such a contrived methodology is informative about people's ability to infer, on a moment by moment basis, other's thoughts and feelings during the rich dynamics of normal social interaction.

A satisfactory measure of empathic accuracy arguably requires researchers to gain access to the content of target's thoughts and feelings as they occur within the context of social interaction. However, measuring people's thoughts and feelings as they occur naturalistically during an interaction is not easy. Not surprisingly, therefore, researchers who have attempted to measure on-line empathic judgements or cognition have encountered methodological difficulties (see Fletcher & Kininmonth, 1991, for a more detailed review). The procedures and instructions in such techniques employed to measure empathic accuracy have typically lacked standardisation. One approach gaining currency has used videotape-review procedures, in which participants review their videotaped interactions and report on their private thoughts and feelings. However, when using such videotape-review procedures, researchers have often incorporated long delays between the original recording of the interactions and the subsequent reviewing sessions. Such lengthy delays may unwittingly encourage subjects to invent cognition and emotions that never occurred in the original discussions. Finally, researchers have frequently asked participants direct questions concerning specific thoughts and feelings, thereby almost certainly priming them to produce cognitions that might otherwise have not been spontaneously generated.

One promising methodology that helps to overcome these problems is the procedure developed by Ickes, Stinson, Bissonnette, and Garcia (1990), in which strangers, independently and immediately, review videotapes of prior dyadic interactions in which they have just participated. In this procedure, participants are required to pause the videotape of their interaction at the points at which they can recall having experienced specific thoughts and feelings, and to make a written, time-logged listing of all such thoughts and feelings. They are then required to watch the tape again; however, on this

pass, the tape is stopped at those points at which their interaction partners had reported particular thoughts and feelings. The participant's task on each of these tape stops is to infer the actual content of their partner's thoughts and feelings. Empathic accuracy is subsequently measured by having independent coders rate the degree to which the judge's inferences match the target's thoughts and feelings.

With two major modifications, Thomas, Fletcher, and Lange (1997) successfully adapted this technique to measure the empathic accuracy of married couples during their discussions of serious relationship problems. First, instead of being unobtrusively videotaped while they were ostensibly waiting for the experiment to begin, the couples knew in advance that they were being videotaped. Previous research by Fletcher and Fitness (1990) has found that this variant of Ickes et al. (1990) video-review technique offers a viable means of examining cognitions as they occur naturalistically in the context of close relationship interaction. Second, when participants were viewing the tape for the second time, they were required, in addition to inferring the partner's thoughts and feelings, to reveal their own thoughts and feelings they had experienced at the same points on the tape. This additional assessment provided a way of measuring, and subsequently controlling for, the degree to which judges assumed that their own thoughts and feelings were similar to the target's thoughts and feelings. Judge's use of assumed similarity is the only Cronbachian confound that seems likely to contaminate empathic accuracy scores (see Thomas & Fletcher, 1997).

General Aims of this Thesis

The study of interpersonal accuracy has flourished over the last decade. The rise of accuracy research has been dominated by studies on the accuracy of personality judgements, with the study of empathic accuracy still in its infancy. Fortunately, both lines of research have moved beyond the issue of simply to what extent people make accurate judgements. We can safely assume that social judgements will sometimes be accurate and sometimes inaccurate. The new wave of accuracy research has largely focused on the moderators of accuracy. That is, what are the features that enhance or diminish accurate judgement?

Although this research on moderators has been fruitful, there are a number of important, hitherto unanswered questions that need to be addressed, which this thesis grapples with. Specifically, the current research breaks new ground in four major ways. First, it provides a comprehensive examination of the influence of the perceiver-target relationship (otherwise referred to as “good information” or the “good relationship” or the “acquaintanceship effect”) on judgement accuracy. In short, this thesis addresses the question of the extent to which increased acquaintanceship is associated with greater knowledge and understanding of others. This seemingly obvious question has in fact engendered considerable controversy in social and personality psychology over the last thirty years. A related goal of the current research was to determine the characteristics of the good relationship. That is, what are the relationship-level variables that predict empathic accuracy attained by well-acquainted judges?

The second major aim was to investigate the existence of individual differences in empathic ability at various levels of acquaintanceship. Although the concept of the good judge of personality has tantalised accuracy researchers, inconsistent results combined with severe methodological problems, led many to abandon the search entirely. In fact, some researchers have asserted that individual differences in accuracy do not exist, or at least are too limited to be important (e.g., Kenny, 1994). It is true that a few methodologically rigorous studies have recently found evidence for a substantive individual difference component in interpersonal accuracy, for both empathic and trait judgements (e.g., Marangoni, Garcia, Ickes, and Teng, 1985; Vogt & Colvin, 1998). However, an important issue previously not dealt with by such research is the extent to which individual differences generalise across a broad range of judge-target relationships. For example, is the good judge of strangers also a good judge of his or her friend or dating partner? To fill this lacuna, this thesis examines the evidence of the good judge across different levels of acquaintanceship in the domain of empathic accuracy. An associated goal was to determine the characteristics of the good judge. That is, a number of theoretically guided individual difference variables were tested to determine whether they distinguished between consistently good and poor performance.

A third important aim of the current research was to investigate the existence of individual differences in target’s readability in empathic accuracy. That is, are certain targets’ thoughts and feelings easier to read than others? Prior research has found good

evidence for the good target when judged by relative strangers. However, the current research breaks new ground by examining the moderating effect of target's readability at different levels of acquaintanceship.

The final aim of this thesis was to investigate the nature of the link between trait accuracy and empathic accuracy at different levels of acquaintanceship. That is, whether the good judges of target's thoughts and feelings are also good judges of that person's personality traits. A number of researchers have speculated that the accurate perception of states is a necessary prerequisite to the accurate perception of traits (e.g., Ickes, 1993). However, this research represents the first attempt (to my knowledge) to integrate these two important domains of interpersonal accuracy research.

To achieve all of these aims, three studies were conducted in this thesis. Study 1, the major foci of the thesis, examined the evidence for three moderators of empathic accuracy assessed during close relationship interaction: To what extent is empathic accuracy moderated by (a) the ability of the judge ("good judge"), (b) the readability of the target ("good target"), and (c) the level of acquaintanceship ("good relationship")? Study 2 investigated the moderating influence of trait visibility ("good trait") and the level of acquaintanceship ("good relationship"), along with their unique interaction, on the accuracy of personality judgements. Finally, Study 3 examined the relationship between trait accuracy and empathic accuracy. Moreover, the research was designed with three important features. First, multiple perceivers judged multiple targets. Second, these judgements were made at three levels of acquaintanceship – dating partners, friends, and strangers. Third, all perceivers judged both the target's underlying psychological traits and mental states.

STUDY 1

THE MODERATORS OF EMPATHIC ACCURACY: THE QUEST FOR THE GOOD JUDGE, GOOD TARGET, AND GOOD RELATIONSHIP

The theoretical framework used to guide and organise the research in this study constitutes a synthesis of Funder's (1995) Realistic Accuracy Model (RAM) and Thomas and Fletcher's (1997) social cognitive approach to empathic accuracy. Although RAM deals with trait judgements, it can also be fruitfully applied to empathic inferences. In essence, Funder's (1995) model states that a trait can be accurately judged if *relevant* behavioural cues are made *available* to the perceiver and if the perceiver *detects* and appropriately *utilises* these cues. RAM also proposes four basic moderators that make accuracy more or less likely by virtue of their influence on one or more aspects of the judgement process: the relevance, availability, detection, and utilisation of behavioural cues. Three of these moderators seem pertinent to research on empathic accuracy: the "good judge", the "good target", and "good information" or the "good relationship" (these terms are used interchangeably in this thesis).

RAM provides a valuable theoretical vantagepoint for investigating the moderators of empathic accuracy. However, this model cannot fully account for the moderating influence of the quality of the perceiver-target relationship on empathic accuracy. The reason is that RAM views accuracy as being largely determined by the diagnosticity of the behavioural information emitted by the target rather than in terms of the cognitive processes that occur within the perceiver. The judge is accorded the rather passive role of observing and immediately processing the target's behaviour, rather than relying on pre-existing knowledge structures to help interpret and supplement the incoming behavioural information (Colvin, Vogt, & Ickes, 1997). Regardless of the nature of the relationship between the judge and the target, Funder (1995) suggests that accuracy is predominantly data-driven.

Although a data-driven perspective seems reasonable for explaining empathic accuracy at relatively low levels of acquaintanceship (e.g., Ickes et al. 1990), it is not well equipped to account for the process of accurate judgement in more intimate perceiver-target relationships. From a social cognition perspective, social judgement is inherently a

theory-driven process (at least in part). People cannot attend to or process the entire welter of data in the social environment. People, therefore, create knowledge structures to cope with the barrage of information. Once established, these cognitive structures constitute a theory-driven influence on the interpretation of incoming data. Moreover, in the context of close relationships, people develop especially rich and enduring theories concerning their partner and their relationship forged over a long history of prior interactions (see Fletcher & Thomas, 1996). Indeed, the research literature dealing with close relationships is replete with demonstrations of the powerful role exerted by pre-existing relationship-specific knowledge structures on relationship judgements (see Baldwin, 1992; Fletcher & Thomas, 1996). In sum, empathic accuracy in close judge-target relationships is likely to be (to an important extent) a theory-driven process.

Thomas and Fletcher (1997) accord an important role for both data-driven and theory-driven processes in determining empathic accuracy. According to their perspective, intimates and strangers alike rely (at least to some extent) upon behavioural data when generating empathic judgements. However, as relationships become closer, rich pre-existing theories of the partner and the relationship are likely to increasingly direct attention to particular types of behaviour, guide the way this data is processed, and influence judgements concerning the target's thoughts and feelings. Therefore, perceiver's reliance on theory vs. data should vary according to the level of acquaintanceship, with well acquainted judges using prior knowledge structures to interpret and supplement incoming behavioural information to a greater extent than less acquainted judges. I will now use this brief theoretical overview as a springboard to mount a detailed discussion of the moderators of empathic accuracy.

The Good Relationship: The Effect of Acquaintanceship on Empathic Accuracy

Previous Research

The common-sense belief that knowing a person will confer advantages in the accuracy of social judgement has created considerable debate in the accuracy literature. On the one hand, advocates of the ecological approach to social perception have asserted that, for the most part, information available in the immediate environment is sufficient for accurate judgement, without the need for extensive cognitive processing (e.g.,

McArthur & Baron, 1983). This approach focuses on the impressive ability of people to detect subtle cues afforded by the target's physical characteristics and behaviour. This view has been bolstered by a line of research that documents accurate judgements of emotion, deception, interpersonal expectancies, and certain personality traits, based on very thin slices of behaviour (e.g., 30-s interactions or even a photograph). In fact, a meta-analysis of this research found that the effect size of .39 for predictive accuracy was higher than many effect sizes of well-established phenomenon in social and personality psychology (Ambady & Rosenthal, 1992). More recently, some psychologists in the close relationship arena have argued (with empirical support) that love is blind, with partners driven more by the need to protect or enhance their relationships than by the desire to be accurate and objective in their judgements (Murray & Holmes, 1996; Simpson, Ickes, & Blackstone, 1995). Such research and theorising has lead some researchers to conclude that accuracy plateaus at very low levels of acquaintanceship - additional information is merely redundant, or even counterproductive (e.g., Wilson & Schooler, 1991).

On the other hand, several accuracy researchers, particularly in the domain of personality, have defended the common-sense view of acquaintanceship. According to RAM, regardless of differences in judges' ability or targets' readability, increased quantity (availability) and quality (relevance) of information should generally lead to greater accuracy of personality judgement (Funder, 1995). In general, personality researchers have accumulated compelling evidence for an acquaintanceship effect in cross-sectional studies, although the picture is not so clear-cut when accuracy is measured longitudinally or when the criterion is interjudge consensus. These issues will be revisited in more detail in Study 2.

The effect of acquaintanceship on empathic accuracy has received scant attention. Moreover, the results of relevant studies that have used the Ickes et al. (1990) tape-review procedure, previously described, have been inconclusive. Stinson and Ickes (1992) provided the first demonstration of the acquaintanceship effect by showing that male friends were more accurate than male strangers when inferring each other's thoughts and feelings. Subsequent research by Graham (1994) found that this moderating effect of good information on empathic accuracy extended to the performance of female friends vis-à-vis female strangers. Both of these studies, however, employed a simple dyadic

design, in which each perceiver judged only one target. One problem with this particular design is that it fails to cleanly separate the acquaintanceship variable from person-based moderators of empathic accuracy; namely, judges' empathic ability and targets' readability (see Funder & Colvin, 1988).

Two recent studies have employed research designs that obviate the problem outlined above. First, Marangoni et al. (1995) used an experimental design to more precisely measure the influence of information quantity on the accuracy of stranger's empathic judgements while watching 30-min videotapes of three client-therapist interactions. Empathic accuracy improved with increasing exposure to two of these targets (or clients). This effect was not invariant, however, as one target remained enigmatic despite the availability of more information. Second, Hancock and Ickes (1996) measured perceivers' empathic accuracy while observing a 6-min generic "getting acquainted" interaction between two complete strangers (targets). Acquaintanceship was treated as a within-groups variable because each perceiver was a friend of one interactant but a stranger of the other. The results showed that in the context of highly scripted and superficial interactions the effect of acquaintanceship was trivial.

Research Aims

Given the methodological limitations and equivocal results of previous research, the current study had three main aims with respect to assessing the moderating effect of good information (or the good relationship). First, I sought to provide a clear demonstration of the effect of the relationship-based moderator on empathic accuracy, by either eliminating or controlling for the possible influence of individual differences in the judges' empathic ability and targets' readability. The second aim was to adopt a fine-grained examination of the moderating effect of relationship level by comparing the empathic accuracy of dating partners, friends of dating partners, and strangers. To date, researchers have dealt with the perceiver-target relationship in a simple dichotomous fashion comparing friends with strangers. The final aim was to test the acquaintanceship effect in the context of observations of content-rich and complex dyadic interactions (i.e., a problem-solving interaction) – contexts in which well-acquainted judges should have the opportunity to make use of their extensive pre-existing knowledge of the targets.

To achieve these aims, the level of acquaintanceship (dating partner vs. friend vs. stranger) was treated as a between-groups variable. The empathic accuracy criterion consisted of the self-reported thoughts and feelings of dating partners that occurred during the course of their videotaped problem-solving interaction. Empathic accuracy was assessed by each type of acquaintance watching the videotaped interaction, and then inferring the partner's thoughts and feelings. Hence, the male and female dating partners were the relationship interactants, with each dating partner serving as both a perceiver and as a target. In contrast, the nominated friend of the dating partners and the stranger served only as perceivers that observed the dating couple's interaction.

Insider versus Outsider Information and the Role of the Relationship

A useful way of conceptualising the effect of information in this type of research design is in terms of the commonly made distinction between an *outsider's* and an *insider's* perspective of an interaction (see Thomas & Fletcher, 1997). An outsider is an observer who merely watches the dyadic interaction, whereas an insider is classed as one of the participants in the close relationship. The category of outsider can be further demarcated into subjective and objective outsiders. *Subjective outsiders* are friends or acquaintances who have at least some degree of pre-interaction knowledge about the relationship interactants, whereas *objective outsiders* are strangers who have no such prior knowledge.

There is good evidence that outsiders and insiders forge divergent interpretations of the meaning of behaviour in romantic relationship interactions (see Surra & Ridley, 1991). This is perhaps not surprising given the different quantity and quality of information available to each type of acquaintance. Objective outsiders have access only to information, derived from the brief interaction that they observe, comprising the verbal and nonverbal behaviour evinced by the participants. When interpreting participants' behaviour, objective outsiders will be forced to rely on normative or culturally shared rules about the meaning of such data. In addition, objective observers may make theory-driven inferences by invoking their stereotypes or general relationship theories (e.g., general beliefs concerning close relationships, gender roles, or physical attractiveness). Judgements based on such limited information can nonetheless be surprisingly accurate. Indeed, previous research carried out by Ickes and his colleagues suggests that the best

predictor of strangers' empathic accuracy is the quantity and quality (i.e. diagnosticity) of behavioural interaction – the more good information available the more accurate are the empathic judgements of strangers (Hancock & Ickes, 1996; Ickes et al., 1990; Marangoni et al., 1995; Stinson & Ickes, 1992).

In contrast to objective outsiders, insiders not only have access to the interactional data and stereotypical information, but they also have at their disposal theories and knowledge concerning their actual partners (the targets) and their specific relationship. Recent research suggests that insiders do not rely on their partner's interactional behaviour in any straightforward way, as a primary source of their empathic judgements. Instead, prior relationship knowledge and theories appear to have a major influence on empathic judgements (see Simpson et al., 1995; Stinson & Ickes, 1992; Thomas et al., 1997). Of course, this conclusion will come as no surprise to those in social cognitive circles. According to a social cognitive approach, lay relationship-specific theories will direct attention to certain types of behavioural information, guide the way this information is processed, and influence how it is used to generate empathic inferences (Baldwin, 1992; Thomas & Fletcher, 1997). Like insiders, subjective outsiders also have access to both observational data and pre-existing theories about specific relationships. Such theories, however, are more likely to constitute general knowledge of the target (personality, attitudes, and so forth) rather than extensive information specifically about the target's dating relationship.

In effect, determining the effect of acquaintanceship on empathic accuracy involves pitting the more theory-driven judgements of dating partners, and to a lesser extent friends, against the more data-driven judgements of strangers. Certainly, data-driven judgements based on thin slices of behaviour can lead to impressively accurate judgements on some dimensions, and high levels of information can sometimes overwhelm or even mislead some judges (Tetlock & Boettger, 1989). However, theory-driven empathic judgements should generally be more accurate in the context of relationship interaction in which specific problems are being discussed. Consider the sources of experience and knowledge that the insiders have access to. First, insiders know one another well in some general sense. Second, they will often have had similar sorts of problem-solving discussions in the past. Third, they may have previously discussed their thoughts and feelings with one another about the specific problems under

discussion. Subjective insiders (friends) will not typically have access to quite the same level of detailed knowledge as the insiders (partners) but they will have normally observed the couple interacting, and will also have sometimes discussed the relationship with one or both of the participants.

In summary, it is predicted that increased acquaintanceship would be related to greater empathic accuracy. Specifically, dating partners would be more accurate than friends, and friends would be more accurate than strangers.

The Good Target: Individual Differences in Target's Readability

Previous Research

The idea of the quintessential good target is thought to be independent of the influence of judge-based and relationship-based moderators. In other words, perceivers should consistently find some targets easier to judge than others, regardless of their own respective levels of judgement ability or the kind of relationship they have with the targets. This hypothesis has been bolstered by a few studies showing that perceivers' judgements of personality have exhibited greater accuracy for some friends than for others (e.g., Colvin, 1993a, 1993b).

Evidence for the good target has also been found in the domain of empathic accuracy. Marangoni et al. (1995) found individual differences in the readability of three targets (clients) when judged by outside observers in the context of a therapist-client interaction. Similarly, Hancock and Ickes (1996) demonstrated that during a brief unstructured interaction between two strangers, target's readability accounted for nearly half of the variance of perceiver's empathic accuracy scores. However, no research has investigated individual differences in target's readability during close relationship interaction. Furthermore, scant attention has been directed to the influence of the relationship between the perceiver and the target on differences in target's readability. Previous research has almost exclusively focused on target effects at minimal acquaintance – with strangers (outsiders) making empathic judgements based on brief observations of targets' behaviour. Although, Hancock and Ickes (1996) included a friend (subjective outsider) as one of the judges in their study, because the "getting acquainted" interaction between the targets was highly scripted and superficial, the

possible effect of pre-existent friendship-specific knowledge on targets' readability may have been nullified.

Research Aims

In this study, I extended previous research on the good target of empathic judgements in two important ways. First, the moderating effect of target's readability was examined at different levels of acquaintanceship. The dating partners who served as targets were judged by three perceivers: the other dating partner (insider), a friend of the dating couple (subjective outsider), and a stranger (objective outsider). Target's readability was assessed by the level of consensus or consistency in empathic performance (as measured by intra-class correlation) attained by the three perceivers when judging the same target. Such a design is a rigorous test of the good target hypothesis, because judge-based and relationship-based moderators could not explain the existence of consensus among judges.

The Role of Behavioural Diagnosticity

Based on the logic of RAM, for the good target of empathic judgements to be found, across the full range of acquaintanceship, two conditions need to be met (Funder, 1995). First, some targets must emit behavioural information during their interaction that is more diagnostic of their inner thoughts and feelings than others. Second, perceivers at each level of acquaintanceship must utilise the available behavioural information when making their empathic judgements. What is the evidence for these two conditions?

Using a variant of Ickes et al. (1990) empathic accuracy procedure, Simpson et al. (1995) developed a measure of behavioural diagnosticity during couples' interactions. Coders viewed each couple's videotaped discussion and stopped the tape at the points at which each partner reported having a thought or feeling. They then read the content of each partner's thought/feeling statements and rated, on the basis of the verbal and nonverbal information conveyed before the thought and feeling was reported, how difficult it would be for an observer to accurately infer the content of participants' reported cognitions and emotions. In this way, Simpson et al. directly measured how openly and directly the couples communicated their thoughts and feelings. They found

that some dating couples' behaviour was more diagnostic of their inner thoughts and feelings than other couples.

Simpson et al's (1995) findings have been bolstered by several studies that have found considerable variability in the extent to which people openly and directly communicate their thoughts and feelings during dyadic interactions (see Thomas & Fletcher, 1997 and Fletcher, Thomas, & Durant, in press). For example, Rusbult, Verette, Whitney, Slovik, and Lipkus (1991) demonstrated that individuals who are highly committed to their close relationships report a willingness to inhibit their impulses to reciprocate destructive behaviour, and instead react constructively. Research by Clark, Pataki and Carver (1996) also showed that some people in relationships commonly feign emotions (such as anger, sadness, and happiness) in order to ingratiate, intimidate, or get help from their partners. In summary, the evidence suggests that some individuals censor the expression of their thoughts and feelings, or feign cognitions and affect more than others.

Even if targets do routinely provide clear and coherent behavioural information concerning their thoughts and feelings, this does not mean judges routinely use this information to derive their empathic inferences. As previously discussed, objective outsiders' judgements should be primarily data-driven, as evidenced by the substantial link between the greater availability of diagnostic information and more accurate empathic inferences with stranger judges (e.g., Hancock & Ickes, 1996; Marangoni et al., 1995). In contrast, Thomas et al. (1997) found no relation between behavioural diagnosticity and empathic accuracy for insider judges (marital partners). Hence, insiders do not seem to base their judgements primarily on the face-value information accorded by the behaviour, even when it is highly diagnostic. Instead, Thomas and Fletcher (1997) argue that elaborate theories held by the partners concerning the relationship are relied upon to help interpret and supplement the incoming behavioural information. Although, previous research has not examined subjective outsiders' reliance on theory versus data during close relationship interaction, it is a reasonable supposition that friends' judgements would also be more theory-driven than strangers' judgements.

To summarise, it was expected that individual differences in target's readability across different levels of acquaintanceship (the "good target") would not be found. Moreover, to test whether differential reliance on data versus theory in making

judgements might account for this null hypothesis, perceivers' empathic accuracy scores at each level of acquaintanceship were correlated with the diagnosticity of target's behaviour. It was predicted that as acquaintanceship increased the perceiver's use of behavioural information would decrease.

The Good Judge: Individual Differences in Empathic Ability

Previous Research

Early research on interpersonal accuracy was dominated by the quest for the good judge of personality and their associated characteristics (see Taft, 1955). Yet, almost fifty years on, the jury is still out over whether certain individuals possess this quintessential ability. On the one hand, the notion of the good judge has occupied an important place in recent theories of accuracy. Advocates of the ecological approach to social perception assert that people exhibit individual differences in attunement to their social environment, mainly for motivational reasons (e.g., McArthur & Baron, 1983). In addition, RAM (Funder, 1995) proposes that good judges must possess a greater capacity for detecting and using the available behavioural cues correctly in order to derive their superior judgements of others' personalities. On the other hand, empirical research has yielded such inconsistent and largely negative results that most current researchers have abandoned entirely the study of individual differences in judgement ability (Kenny, 1994; Schneider, Hastorf, & Ellsworth, 1979). However, this extreme reaction is perhaps premature, given the fact that prior studies have been plagued by a number of methodological shortcomings.

As mentioned previously, the earlier generation of accuracy research was susceptible to the influence of a number of methodological artifacts (Cronbach, 1955). Cronbachian problems aside, a major flaw of even contemporary studies is that they have almost exclusively used simple dyadic designs, in which each perceiver judges only one target. This design is unable to cleanly separate the effect of the judge from that of the target and the judge-target relationship. Hence, such results are inconclusive regarding differences in judgement ability (Kenny, 1994).

The requisite design for studying the good judge is one in which each perceiver judges multiple targets. Consistency of performance across several targets by the same

judge is then used as the defining feature of judgement ability (Colvin, et al., 1997). The few researchers who have used multiple targets have found encouraging results for the good judge of personality (Vogt & Colvin, 1998) and dyadic rapport (Bernieri, Gillis, Davis, & Grahe, 1996).

Within the domain of empathic accuracy, the first study to examine individual differences in empathic ability found substantial cross-target consistency in perceiver's performance when judging the thoughts and feelings of three strangers (clients) in the context of a therapist-client interaction (Marangoni et al., 1995). However, attempts to replicate this finding when judging strangers' thoughts and feelings during mundane "getting acquainted" interactions have generally failed (e.g., Hancock & Ickes, 1996). A likely reason for the discrepancy between these two sets of findings is the different demands placed on perceivers during the two types of judgement tasks. Judging the content of clients' thoughts and feelings during a complex and affect-laden therapeutic interaction is probably a more demanding task for perceivers than making such judgements during a highly scripted and superficial interaction between previously unacquainted strangers, most of whom have natural and therefore quite desultory conversations. Previous research on accuracy has shown that difficult tasks discriminate more effectively between judges than easier tasks (e.g., Fletcher, Rhodes, Rosanowski, & Lange, 1992). To illustrate this point, a very easy exam is less likely to discriminate between good and poor students than a difficult exam. Similarly, mundane "getting acquainted" interactions probably fail to provide the good judge with sufficient opportunity to demonstrate their superior empathic ability, compared to the behaviourally rich and complex interactions between people who know one another.

Research Aims and Predictions

The present research extended the few previous studies on individual differences in empathic ability in three major ways. First, participants made empathic judgements in the context of targets' discussions of their most serious romantic relationship problems. It was predicted that individual differences in the ability of the judge would emerge more strongly than in previous research, because of the subtlety and high difficulty level of the empathic accuracy task used in this study. Second, because of the variability in the level of acquaintanceship, I was able to assess the consistency of empathic performance across

quite different perceiver-target relationships in which: (a) targets were all strangers, (b) two targets were friends and the other targets were strangers, and (c) one target was a dating partner and the other targets were strangers. Such a design is a rigorous test of “true” individual differences in empathic ability, because the existence of moderate to strong correlations in all cases cannot be explained in terms of target-based and relationship-based moderators. I made the straightforward prediction that individual differences in empathic ability would be found, regardless of the type of judge-target relationship.

Finally, another contribution of this study was to search for the individual characteristics that might differentiate consistently accurate judges from consistently inaccurate ones. In the absence of prior relevant research, I used pertinent theories such as RAM (Funder, 1995), along with a recent meta-analysis of the correlates of interpersonal accuracy (Davis & Kraus, 1997), to guide my specific hypotheses. Note, however, that for the purposes of this study, the findings of the meta-analysis can only be suggestive, because it encompasses a wide range of accuracy judgements typically derived from simple dyadic designs.

The Characteristics of the Good Judge

Verbal Intelligence

Intellectual ability has long been regarded as a predictor of accuracy in social judgement. Two comprehensive reviews of early research cautiously implicated intelligence as one of the few psychological properties of the good judge (Bruner & Tagiuri, 1954; Taft, 1955). Similarly, Davis and Kraus (1997) in their meta-analysis of 36 post-Cronbachian studies, concluded that intelligence contributes in a reliable, yet modest fashion to judgement accuracy (with a mean effect size of .23). According to RAM, (Funder, 1995) greater IQ could plausibly improve both the likelihood of judges detecting relevant cues from the large amount of behaviour evinced by the target, and the subsequent use of this information in a valid manner. Based on such research and theorising, it was predicted that higher levels of verbal intelligence would be associated with greater levels of empathic ability.

Attributional Complexity

Another individual difference variable implicated by RAM (Funder, 1995) that might influence judgement accuracy is attributional complexity. Attributional complexity refers to judges' propensity to make complex causal explanations for social behaviour (see Fletcher, Danilovics, Fernandez, Peterson, & Reeder, 1986). Judges with a preference for generating more elaborate explanatory theories for behaviour could plausibly be particularly sensitive to diagnostic behavioural cues and be especially capable of effectively integrating these cues to derive judgements that are more accurate. Davis and Kraus' (1997) meta-analysis demonstrated that measures of cognitive complexity were the best predictors of judgement accuracy with a mean effect size of .27.

Previous research (in non-relationship domains) with the Attributional Complexity Questionnaire have shown that high levels of complexity predict more accuracy in trait attributions under two conditions – high motivation and high difficulty level of the task (Fletcher et al., 1992; Fletcher, Reeder, & Bull, 1990). Both of these conditions are present in the current research. Moreover, although it is a self-report measure, it does not directly assess self-reports of empathy or accuracy. Instead it (reliably and validly) assesses complexity in attributional schemata. Moreover, the evidence indicates that complexity of social schemata is a promising marker of social intelligence. In summary, it was hypothesised that higher levels of attributional complexity would be associated with increased accuracy of empathic judgements.

Self-Reported Dispositional Empathy

Despite intensive research efforts over the past fifty years, self-report measures of empathic ability have proved to be particularly poor predictors of actual empathic accuracy (see Ickes, 1993). Davis and Kraus' (1997) meta-analysis revealed a mean effect size of nearly zero for various measures of self-reported dispositional empathy. Moreover, Mortimer (1996) found that even when self-reported accuracy ratings were measured on an inference-by-inference basis, between 85 to 90% of judges lacked even a minimal degree of accurate metaknowledge concerning their own level of empathic accuracy. According to Ickes (1993), a likely reason for this lack of metaknowledge is that perceivers rarely seek explicit feedback on the accuracy of their judgements, and on the infrequent occasions such feedback is provided it is often misleading or

uninformative regarding the perceiver's *relative* level of empathic accuracy. Thus, it was predicted that neither judges' ratings of their general perspective-taking ability, nor their self-reported assessments of their capacity to take their partner's perspective regarding the particular problems discussed, would predict their actual levels of empathic accuracy.

Gender

Claims concerning the superiority of women's intuition, compared to men's, abound in popular culture. However, a number of reviews of the accuracy literature have cast doubt over this conventional wisdom (e.g., Eisenberg & Lennon, 1983; Graham & Ickes, 1997; Lennon & Eisenberg, 1987). Although women routinely report greater self-perceived empathic ability and considerably more emotional sensitivity than men, more direct measures of empathic accuracy have often failed to substantiate these claims. According to Graham and Ickes (1997), the few studies that have demonstrated traditional gender differences in empathic accuracy contain a common thread – they all involved judgement tasks that seemingly motivated female perceivers to attain higher accuracy than males (but c/f Thomas & Fletcher, 1997).

One domain where women appear to have greater motivation (and knowledge) than men is close romantic relationships. A number of theorists have proposed that a woman's self concept is more relational and interdependent than a man's self concept (Bakan, 1966; Cross & Madson, 1997; Markus & Kitayama, 1991). This view has been bolstered by good evidence that women spontaneously focus more attention on relationship information, possess more elaborate and complex relational schemas, and talk more about relationships than men (Acitelli, 1992; Acitelli & Young, 1996; Cross & Madson, 1997). Given their level of general relationship expertise, and the nature of the current research task, it was predicted that women would attain greater levels of empathic accuracy than men.

Big-Five Factors of Personality

The final set of variables that were examined in the present study as potential characteristics of the good judge were the five broad dimensions of personality known as the Big-5: extraversion, agreeableness, conscientiousness, emotional stability, and intellect (Goldberg, 1992). Extraversion and emotional stability have received the most

attention in the accuracy literature. According to RAM (Funder, 1995), the greater amount and type of interpersonal experience gained by extraverts might enable them to make more accurate judgements than introverts. This notion is consistent with Akert & Panter's (1988) finding that extraverts are better judges of nonverbal behaviour in social interaction. In addition, RAM suggests that individuals that are especially anxious or defensive are unlikely to possess high levels of judgement ability. The corollary of this prediction is that the good judge is likely to possess a high degree of emotional stability. However, meta-analytic results have revealed that neither extraversion nor neuroticism is reliably and meaningfully related to interpersonal accuracy (Davis & Kraus, 1997). Given these contradictory findings, no particular hypotheses were made regarding the association between the Big-Five factors of personality and empathic accuracy.

The Characteristics of the Good Relationship

The most straightforward prediction concerning the effect of the relationship between target and judge is one already canvassed; namely, that as acquaintance increases so will accuracy increase. In this section, however, I will discuss some of the characteristics, that might be related to accuracy, that describe the relationships of the insiders (partners) and subjective outsiders (friends).

Shared Cognitive Focus

One relationship-level variable that should influence the empathic accuracy attained by dating partners, consistent with RAM's postulate that behavioural cues must be both detected and utilised, is shared cognitive focus. That is, dating partners who concentrate on the topic at hand, pay close attention to the interpersonal flow of information, and adopt a shared frame of reference and interpersonal cognitive focus should achieve more accurate inferences of their respective partner's thoughts and feelings. The design of this study allowed me to measure shared cognitive focus because at each point of the tape that both dating partners reviewed, self-reports of what each partner was thinking and feeling were available. By rating the degree of similarity between each dating partner's thoughts and feelings, I obtained a measure of shared cognitive focus. In the first study of its kind, Thomas et al. (1997) found a strong

positive link between shared cognitive focus and empathic accuracy in married couples. Hence, it was expected that dating partners who attained higher levels of shared cognitive focus would achieve higher levels of accuracy when judging their respective partner's thoughts and feelings.

Relationship Satisfaction

The possible link between relationship satisfaction and empathic accuracy has engendered considerable debate. One common claim (consistent with conventional wisdom) is that partners' satisfaction with their romantic relationship is positively associated with empathic accuracy. Advocates of this stance assert that unhappy couples function in a negative emotional climate in which uncharitable theory-driven judgements about their partner and their relationship produce the inevitable outcome of mutual misunderstanding (e.g., Noller & Ruzzene, 1991). Initial evidence for this view was derived from clinical observations of the communication patterns of unhappy couples when discussing conflicts in their relationships (see Gottman, 1979). More recently, researchers have used empathy scales (e.g., Davis & Oathout, 1985) and measures of affective sensitivity concerning nonverbal cues to demonstrate the superior accuracy of satisfied couples (e.g., Noller & Vernados, 1986). Alternatively, some researchers have argued that love is blind, that people in happy relationships are motivated to view each other through rose-tinted glasses, whereas those in unhappy relationships are prepared to face a bleaker, but more accurate reality (Hendrick & Hendrick, 1988; Murray & Holmes, 1996). This view is bolstered by research findings that for judgements of various attitudes (Sillars, Pike, Jones, & Murphy, 1984) and traits (Murray & Holmes, 1996), greater accuracy (after controlling for the effects of assumed similarity) was related to lower, rather than higher levels of relationship satisfaction.

One of the difficulties in reconciling these conflicting viewpoints and data is that many of the associated studies are methodologically flawed in at least two ways. First, none of the studies measured empathic accuracy as it occurred on a moment-by-moment basis during close relationship interaction. Second, the studies reporting a positive relation between relationship satisfaction and accuracy have invariably failed to account for the possible influence of assumed similarity on judgement accuracy. There is evidence that partners who are more satisfied with their relationships not only tend to be

more similar, but also to assume that they are more similar to each other than partners who are less satisfied in their relationships (Acitelli, Douvan, & Veroff, 1993; Sillars et al., 1984; Sternberg & Barnes, 1985). Hence, it is plausible that a positive correlation between relationship satisfaction and judgement accuracy is an incidental by-product of actual and assumed similarity. In the only study so far to measure on-line empathic accuracy, while controlling for the effects of assumed similarity, Thomas et al. (1997) found that marital satisfaction was unrelated to empathic accuracy. Given the ambiguity of previous research and theorising, I made no prediction concerning the association between dating partner's relationship satisfaction and their empathic accuracy.

Perceived Relationship Closeness

As argued previously, knowledge structures or theories concerning one's own partner (or friend) and relationship may play an important part in driving empathic judgements (Thomas & Fletcher, 1997). Clearly, however, not all theories are created equal. An important dimension that might influence the validity of such theories is the degree of closeness or intimacy between partners (or friends). One of the defining features of intimate relationships is the extent to which people disclose their innermost thoughts and feelings to each other (McAdams, 1988; Reis & Shaver, 1988). Hence, the greater access to diagnostic information by people in more intimate relationships should enable them to forge more extensive and accurate working models of each other and their relationship. Consistent with this view, Gesn (1995) revealed that closer, as opposed to distant, same-sex friends attained greater levels of accuracy when inferring each other's thoughts and feelings. Thus, it was predicted in the present study that dating partners and friends who had more intimate relationships with the targets would be more accurate in their judgements of the targets' thoughts and feelings.

Problem-Specific Self-Disclosure

The final type of relationship-specific variable examined was problem-specific disclosure. The influence of such information on judges' empathic accuracy scores was assessed for both the confidant (the friend) and the disclosers (the male and female dating partners). I will discuss each in turn.

Close friends are sometimes used as sounding boards for people to vent their raw and uncensored thoughts and feelings about the serious problems in their romantic relationships. The provision of such information is likely to provide the single most useful and veridical knowledge base for making empathic judgements. Thus, I predicted that higher levels of problem-specific disclosure by dating partners should be associated with superior levels of friends' empathic accuracy. However, this link is likely to be moderated by the gender of the discloser, consistent with the prior analysis of gender. There is good evidence that women disclose more personal information than men, particularly in the context of friendship and other close relationships (see Dindia & Allen, 1992 for a meta-analytic review). Moreover, the quality of women's disclosure tends to be more intimate and evaluative, whereas men's disclosure tends to be more superficial and descriptive (see Cross & Madson, 1997). Therefore, it was expected that higher levels of friend's empathic accuracy would be associated with higher levels of female dating partner's (but not male dating partner's) problem-specific disclosure to the friend.

What effect will self-disclosure to the friend have on the empathic accuracy of the discloser? Although claims concerning the cathartic and psychological value of confession and intimate disclosure abound, empirical findings concerning this issue are mixed (Kelly & McKillop, 1996). On the one hand, there is evidence that revealing secrets and disclosing serious personal problems or traumas can produce positive psychological and physical consequences for the discloser, particularly in a non-judgemental therapeutic context (e.g., Pennebaker, Barger, & Tiebout, 1989; Regan & Hill, 1992; Vangelisti, 1994; Wegner & Erber, 1992). On the other hand, some researchers argue that the costs of revealing may actually outweigh the associated benefits, especially if such disclosure elicits a negative, unhelpful, or judgemental reaction from the confidant (see Kelly & McKillop, 1996 for a review). To date, the limited amount of research that has focused on the cognitive benefits for the discloser outside of a therapeutic context, such as gaining fresh insight and perspective on their problems, have invariably used self-reported measures of understanding. The present study provides an important contribution to the literature by *directly* measuring the link between prior self-disclosure to a friend about the discloser's romantic relationship problems and empathic accuracy attained by the discloser during a subsequent discussion of such problems with their dating partner. Based on the evidence that women reveal

higher quantity and quality personal information than men, and in keeping with the previous discussion of gender, the beneficial influence on empathic accuracy was expected to be confined to female disclosers. Thus, I predicted that higher levels of empathic accuracy attained by female dating partners would be associated with higher levels of problem-specific disclosure to the friend provided by female (not male) dating partners.

Overview of Study 1

In this study, I tested for the moderating effects of the good judge, the good target, and good relationship on empathic accuracy during close relationship interaction. I also sought to determine some of the characteristics of the good judge and the good relationship. To achieve these aims multiple perceivers judged multiple targets at different levels of acquaintanceship. Fifty dating couples participated in a ten minute videotaped discussion of two serious relationship problems. Dating partners then independently reviewed the taped interaction twice. On the first occasion, partners provided a written time-logged listing of the thoughts and feelings that they could remember experiencing during the interaction. These times were swapped between partners. On the second occasion, partners stopped the tape at the specified times and described both their own on-line thoughts and feelings and their judgements of their partner's thoughts and feelings. Similarly, fifty nominated friends and fifty strangers observed the dating couple's videotaped interactions and inferred the thoughts and feelings of each partner. Later, dating couples, friends, and strangers made empathic judgements of videotaped married couple's problem-solving interactions. In all cases, the perceivers did not know the married couples. Observer coders later rated perceiver's empathic judgements for accuracy, and also measured shared cognitive focus and assumed similarity of thoughts and feelings for the dating couples.

Summary of Predictions

Good relationship. Empathic accuracy would vary as a function of the type of judge-target relationship. More specifically, dating partners would be more accurate than friends, and friends would be more accurate than strangers.

Good target. Individual differences in target's readability across different levels of acquaintanceship would not be found, based on the prediction that different levels of acquaintanceship would be associated with differential reliance on the surface, stereotypically-informative behavioural information.

Good judge. Individual differences in judges' empathic ability would be found, regardless of the type of judge-target relationship.

Characteristics of the good judge. The good judge across different levels of acquaintanceship would be more likely to be female, and to be characterised by higher levels of verbal intelligence and attributional complexity. However, judges' ability would not be associated with self-reported levels of empathic ability.

Characteristics of the good relationship. Good judges of their own dating partner would be in a more intimate dating relationship and attain higher levels of shared cognitive focus than poor judges. More intimate friends would be better judges of the dating partners' thoughts and feelings than less intimate friends. Higher levels of friend's empathic accuracy would be associated with higher levels of female dating partner's (but not male dating partner's) problem-specific disclosure to the friend. Greater levels of empathic accuracy attained by female dating partners would be associated with greater levels of problem-specific disclosure to the friend provided by female (not male) dating partners.

In addition a number of exploratory issues were raised, such as the relationship-specific link between relationship satisfaction and dating partners' empathic accuracy as well as the general association between the personality characteristics of the judge and the accuracy of their empathic judgements across different levels of acquaintanceship.

Method

Participants

The initial group of participants (hereafter referred to as 'the dating partners or dating couple') comprised 50 couples currently involved in premarital heterosexual relationships that had been dating their partner for at least two months. The mean age of the dating partners was 22.3 years ($SD = 4.9$). Of the total sample, none reported the relationship as casual, 14% reported the relationship as steady dating, 38% described the

relationship as serious dating, and 48% were living together. Relationships had been in progress a mean time of 16.5 months ($SD = 12.47$). Dating partners' mean education level was 4.6 ($SD = .93$) on a scale ranging from 1 = leaving high school without a qualification and 5 = university study or the equivalent.

Each dating couple recruited an individual who knew at least one of the partners well and was willing to participate. Sixty six percent of these individuals described themselves as primarily being a "friend" of the female dating partner, 20% as an "acquaintance", and 14% as a "family member". For the male dating partners, 60% of those recruited described themselves as primarily a "friend", 30% as an "acquaintance", and 6% as a "family member". Given that it was the most common description for both male and female dating partners, these individuals are hereafter referred to as 'the friends'. Of the 50 friends, 33 were female and 17 were male. On average these friends had known the female dating partner for 80.2 months ($SD = 88.5$) and the male dating partner for 46.2 months ($SD = 62.3$). The mean age of the friends was 24.9 years ($SD = 10.51$) and their mean education level was 4.5 ($SD = .99$).

The final group of participants was 25 females and 25 males that did not know the dating partners (hereafter referred to as 'the strangers'). The average age of the strangers was 22.2 years ($SD = 3.7$). Strangers' mean education level was 4.8 ($SD = .61$).

In summary there was a total sample size of 200 (100 dating partners (50 dating couples), 50 friends, and 50 strangers). Each participant was paid \$15 for his or her participation in the research.

Procedure

Session 1: Dating Partners. The dating partners responded to a campus advertisement requesting couples to participate in a relationship interaction study. Partners completed the Sociodemographic, Big-5 Personality (stereotype ratings), General Perspective Taking, and Attributional Complexity scales in their own homes prior to coming to the laboratory. They were requested to complete these scales independently, to refrain from discussing any of the contents with their partners, and on completion to seal them in the separate envelopes provided.

On arrival at the laboratory, dating partners independently completed the Big-5 Personality (self and partner ratings), Inclusion of Other in the Self, and Relationship Satisfaction scales.¹ Partners then listed in order of seriousness two problems in their relationships that they considered to be serious or were currently experiencing conflict over. Two of these problems were then chosen by the experimenter in such a manner that the problems to be used in a later discussion fairly represented the perceptions of both partners. If the first-ranked problems were different, then both of these conflicts were selected as the two discussion topics. If the first-ranked problem was the same for each partner, this issue was selected. If the second-ranked problem also was the same for both partners, this problem was then chosen. However, if the second-ranked problem was different and the first-ranked problem was the same, the second discussion topic was selected at random. The experimenter then wrote down the two problems selected on a prepared sheet, which was left in plain view on the coffee table. Couples then independently completed the problem-specific perspective taking scale for both problems.

The videotaped interaction was set in a pleasantly furnished and spacious soundproof laboratory. Couples sat around a coffee table, with microphones attached to their lapels, while facing one another in two chairs positioned at right angles so that the wall-mounted camera could obtain a clear picture of both partners. An adjacent laboratory housed all the recording equipment (a microphone and intercom system, two connected videocassette recorders and two colour television monitors).

Before discussing their relationship problems, partners participated in a brief debate (based on a procedure developed by Funder & Colvin, 1991). The experimenter handed each partner a clipboard and gave the following instructions:

“The next thing I do is hand each of you a pad of paper because some people like to be able to take some notes during the next part of the experiment. That is because the next part calls for the two of you to have a short debate. Specifically the topic I have people debate is the use of capital punishment, because most people can come up with at least some arguments on both sides of that issue. I’ll just flip a coin and have [name of the subject] call it. If it comes up what you

¹ Various other scales were also completed at this point, the results of which did not form part of this study and, hence, are not reported here.

call, you will be in favour of capital punishment, and if it doesn't you will be against it. [The experimenter then flipped the coin.] Okay, the debate will last about five minutes. I'll just give you a short time to collect your thoughts and then we'll start."

The experimenter left the room and then after a few minutes, activated the video recorder and told the couple to begin the debate. On completion the experimenter re-entered the room and told the couples that they were now going to discuss their two relationship problems. Partners were instructed to forget about the camera and to behave as naturally as possible, as if they were having the conversation in their own homes. Couples were asked to attempt as far as possible to resolve the problem being discussed. They were also assured that the room was soundproof and their tapes and data would be not be shown to anyone else without their permission. At this point, the experimenter left the room to start the video equipment, and couples were instructed to discuss each problem for 5-mins. Two copies of the taped discussions were produced simultaneously, with a running count of the time elapsed electronically embedded in the corner of the screen.

Collection of thought-feeling data: Dating Partners. The procedures used in the next two phases of the research were based on those developed by Ickes and his colleagues (Ickes et al., 1990), Fletcher & Fitness (1990), and Thomas et al. (1997). After completing their discussions, couples were separated and partners were moved to separate soundproof laboratories and informed that they were going to independently review a videotape of their discussion. Partners were instructed that whenever they could recall experiencing a particular thought or feeling, they should (a) stop the tape using the pause button on the remote control; (b) indicate the time elapsed in seconds; (c) write a clear and candid description of the specific thought or feeling they had experienced; and (d) restart the tape.

I stressed to partners that they were to describe only those thoughts and feelings that they could distinctly remember experiencing during the discussion and not to construct new thoughts and feelings. They were also assured that their partners would not gain access to information subsequently provided. A written summary of these instructions remained with each participant. After an initial practice session to become

familiar with the equipment and the procedure, each partner was left to review the tape in privacy.

Collection of empathic accuracy data: Dating Partners. For the second phase of the review procedure, the experimenter collected each partner's written statements and selected a maximum of 10 thoughts and feelings transcribing the times at which these thoughts and feelings occurred. All partners satisfied the criteria of registering at least 3 thoughts and feelings and thus qualified for this second phase. Statements were selected in such a way as to ensure that there were roughly equal numbers of thoughts and feelings for each problem, and that at least a 20-s interval was maintained between each statement. A list of the transcribed times was then exchanged between partners.

Each partner then played the tape again. This time the partners were instructed to pause the tape at each specified time, and then to describe for the interaction immediately preceding that point, explicitly (in writing), (a) what they themselves were thinking or feeling and (b) what they believed their partners were thinking or feeling.

Session 2: Dating Partners. Approximately one week later (between five and nine days), dating partners returned to the laboratory and independently viewed a videotape, generated during an earlier study by Thomas et al. (1997), of two married couples discussing a relationship problem for five minutes each.² The experimenter showed the dating partners a few seconds of each taped discussion, and then asked whether he or she had ever seen the married couple before. No cases of prior acquaintanceship with any of the targets were reported.

After a brief description of the nature of each discussion problem, the dating partners were instructed to watch the first discussion, pausing the tape at each of the 3 specified points at which the wife had previously reported having had a thought or feeling, and infer (in writing) the content of those thoughts and feelings. Dating partners then watched the first discussion again, and inferred the content of the husband's three thoughts and feelings. The same procedure was repeated for the second 5-min discussion, but on this occasion the husband's three thoughts and feelings were inferred prior to the wife's three thoughts and feelings. After completing this task, partners were given 15 minutes to complete a Verbal Intelligence test. Finally, dating partners were debriefed and thanked for their participation in the study. They were also asked to give

their written consent to permit a nominated friend and a stranger to view their videotaped problem-solving discussion. All partners complied with this request and gave their assurance that they would not talk to the friend about any of the details of the study.

Sessions 1 & 2: Friends. From two to three weeks later, friends completed the Sociodemographic, Big-5 Personality (male and female stereotype version), General Perspective Taking, and Attributional Complexity scales in their own homes. On arrival at the laboratory, friends also completed the Inclusion of Other in the Self scale for each dating partner, rated themselves and both dating partners on the Big-5 personality scale, and then rated the extent to which they had previously talked with each dating partner about the two nominated relationship problems.

Following a detailed explanation by the experimenter of the next (empathic inference) phase of the study, friends viewed the dating couple's videotaped 10-minute interaction twice. The order of inferring the male and female dating partner's thoughts and feelings was counterbalanced. On the first viewing, friends paused the tape at those points where the first designated dating partner had reported particular thoughts and feelings, and then inferred the content of those thoughts and feelings. Next, friends played the tape again, but on this occasion, they inferred the thoughts and feelings of the other dating partner. About one week later, each friend returned for Session 2. Friends then inferred the thoughts and feelings of the married couples and completed the Verbal Intelligence test using the same procedure as the dating partners (see previous description).

Sessions 1 & 2: Strangers. The strangers responded to a campus advertisement requesting individuals to participate in a study on person perception. At the beginning of each session, the experimenter ensured that the strangers did not know either the randomly selected dating couple or the married couples on the videotapes. The procedure for the strangers in both sessions of this study was identical to that of the friends when completing questionnaires and making empathic inferences of both the dating and married couples. However, there were two exceptions. First, strangers did not complete any scales concerning their relationship with the dating partners (as they had none). Second, prior to inferring the thoughts and feelings of the dating couple, they watched a videotape of that same couple's debate over the use of capital punishment. It

² Prior permission was obtained from these marital couples to use their tapes as stimulus material.

was ensured that both the friend and the paired stranger inferred the thoughts and feelings of the male and female dating partner in the same order as regards sex of partner. The experimenter explained that partners had been randomly assigned to a particular side of the debate, and therefore the views expressed did not necessarily represent their true attitudes. Strangers then rated each dating partner on the Big-5 Personality scale. The order of judging the male and female dating partner’s personality was counterbalanced.

Design

The design of perceiver and target relationships for Session 1 of this study is depicted in Figure 1. The level of acquaintanceship (dating partner vs. friend vs. stranger) was treated as a between groups variable. The male and female dating partners were the active participants, with each dating partner serving as both a perceiver and as a target. In contrast, friends and strangers served only as perceivers that observed the dating couples’ interaction.

PERCEIVER:

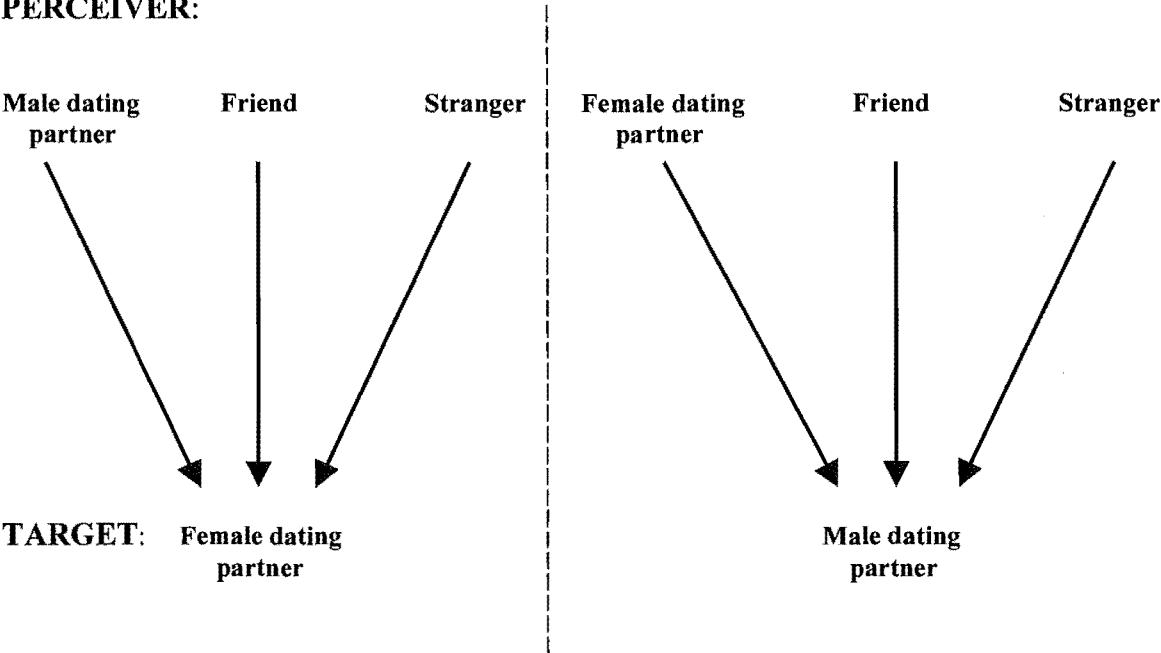


Figure 1. A diagrammatic representation of the perceiver and target relationships in phase 1 of the study.

Figure 2 portrays the design of the perceiver-target relationships for Session 2. On this occasion, the three groups of participants acted solely as perceivers that observed the married couples' interactions. The husbands and wives were the active participants and served only as targets. In all cases, the perceivers and targets were strangers.

PERCEIVER:

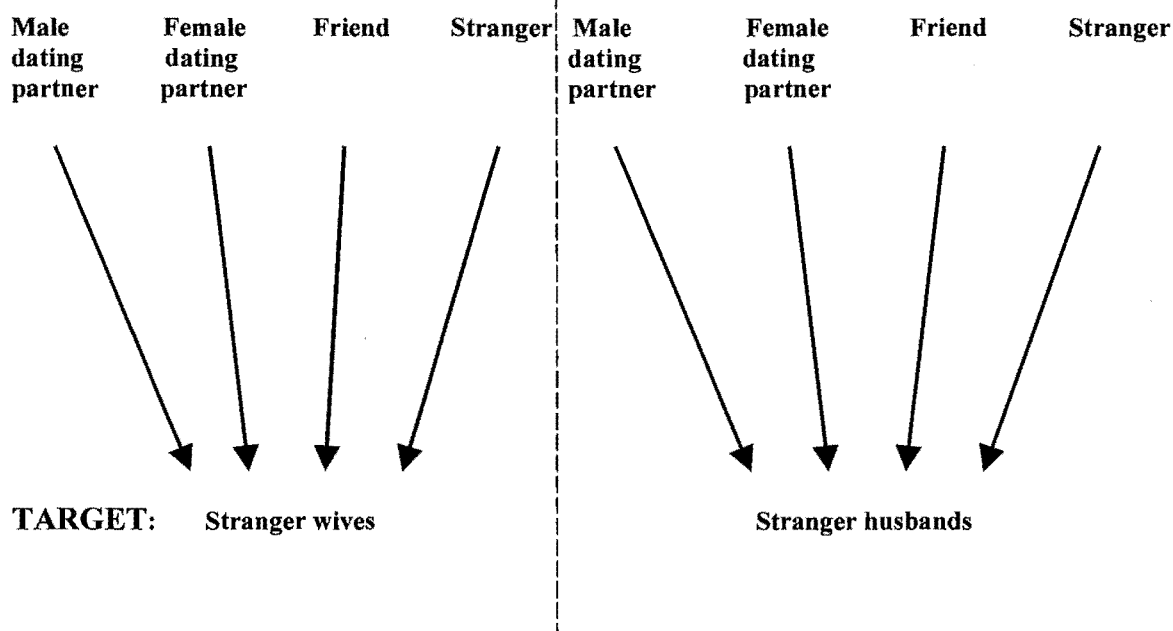


Figure 2. A diagrammatic representation of the perceiver and target relationships in session 2 of the study.

Measures Derived From Observer Coding

Observer ratings of shared cognitive focus, assumed similarity, and empathic accuracy. Two coders estimated the degree of similarity between perceivers' and targets' statements by independently examining the taped discussions in conjunction with the thought and feeling protocols. A 3-point scale was used, where 1 = *essentially different content*, 2 = *somewhat similar, but not the same, content*, and 3 = *essentially the same content* (see Ickes et al., 1990; Thomas et al., 1997). Coders' comparisons of the thought-feeling data yielded three separate measures: shared cognitive focus, assumed similarity, and empathic accuracy. The calculation of these measures is depicted in Figure 3, from the point of view of each dating partner in a dyad.

The shared cognitive focus score was based on the actual similarity between the content of the thoughts and feelings at the same point on the tape. Coders calculated the score by considering the degree to which the same topic was being addressed, ignoring whether the same partner was being referred to or whether the thoughts and feelings were in agreement. For example, if the woman reported that she was unhappy with the division of household labour, and the man reported that he was happy with the division of household labour, these statements would have obtained the maximum shared cognitive focus score of 3.

The assumed similarity score indicated the similarity between the content of each partner's self-reported thoughts and feelings and his or her inference regarding the partner's thoughts and feelings. The criteria here were similar to those used for shared cognitive focus.

The empathic accuracy score characterised the extent to which each target's self-attributed thoughts and feelings matched the content of the perceiver's inference. The associated rating used more stringent criteria than the previous two ratings: Partners needed to be accurate in terms of both target (male or female) and content. For example, to obtain the maximum score of 3 for the female's thought that she was unhappy with the division of household labour, then the male needed to state this; a statement that she was *happy* with the division of labour would have received a score of 2.

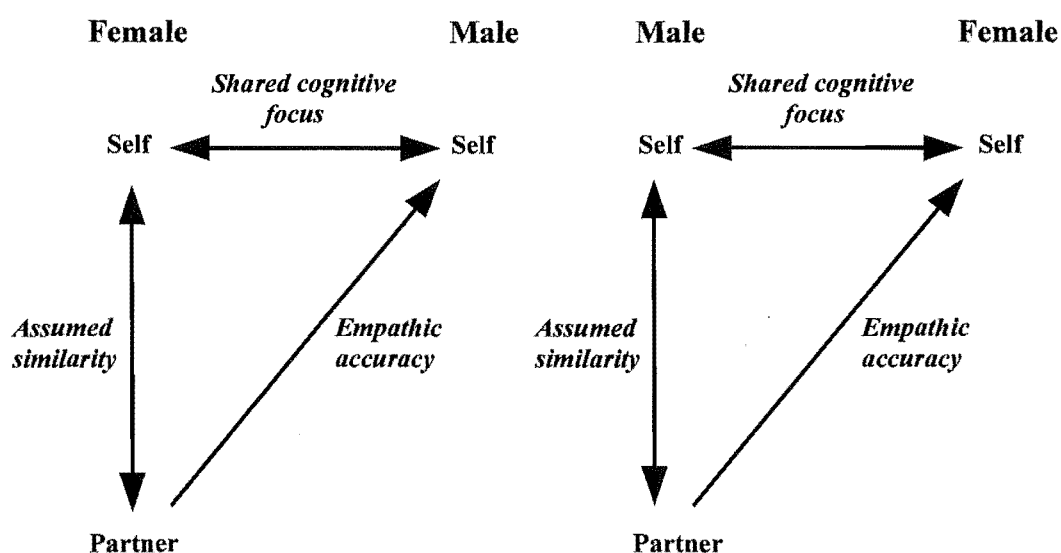


Figure 3. A diagrammatic representation of the rater similarity judgements derived from the thought/feeling statements.

Coders' judgements yielded adequate inter-rater reliabilities: .74 for shared cognitive focus, .75 for assumed similarity, .79 for dating partner empathic accuracy, .79 for friend empathic accuracy, and .76 for stranger empathic accuracy. The inter-rater reliability for empathic accuracy of the stranger married couples (with all three groups of perceivers combined) was .87. All disagreements were resolved by discussion and while re-examining the taped interactions. The final ratings for each perceiver were summed and divided by the number of inferences made to provide mean summary scores for each construct. These ratings were then converted to percentages for ease of readability.

Observer ratings of behavioural diagnosticity. Two coders independently read each target's self thought-feeling statements used in the empathic accuracy coding task, in conjunction with viewing each dating couple's videotaped discussion, stopping the tape at the points indicated. On the basis of the verbal and nonverbal information conveyed in the 30-s of the interaction prior to the reported occurrence of each thought or feeling, coders rated how difficult it would have been for an observer to accurately infer the content of participant's thoughts and feelings. These ratings were made on a 3-point scale, where 1 = *relatively difficult to accurately infer the thought-feeling*, 2 = *neither particularly easy nor difficult to accurately infer the thought-feeling*, and 3 = *relatively easy to infer the thought-feeling* (for more details, see Simpson, et al., 1995; Thomas et al., 1997).

Raters achieved a high level of interrater reliability across all entries ($r = .87$). The few disagreements that occurred were resolved by discussion. These final ratings were tallied for each target and divided by the total number of thought-feeling entries analysed to produce mean behavioural diagnosticity ratings.

Self-Report and Individual Difference Scales

Attributional complexity. This questionnaire consisted of 28 items using 7-point Likert scales (endpoints: strongly disagree to strongly agree) that assessed the complexity of participant's attributional schemata for human behaviour (Fletcher et al., 1986). In previous research it has demonstrated good reliability, convergent validity, and predictive validity. In this study, an internal reliability coefficient of .93 was attained.

Relationship satisfaction. This six-item questionnaire comprised four global judgements on 7-point Likert scales that measure perceptions of relationship happiness, general relationship satisfaction, relationship stability, and the seriousness of relationship problems. This scale has shown good reliability and validity in previous research (Fletcher & Fitness, 1990; Fletcher, Fitness & Blampied, 1990). It attained an overall internal reliability coefficient of .88 in this study.

Relationship closeness. Dating partners described their romantic relationship, and friends described their friendship, on the Inclusion of the Other in the Self (IOS) scale (Aron, Aron, & Smollan, 1992). This single item scale consisted of a series of overlapping circles, ranging from no self-other overlap to extensive self-other overlap, from which participants selected the picture that best described a particular relationship. The IOS scale measures both feelings of closeness and behaviours associated with closeness, and has demonstrated good reliability and validity (Aron et al., 1992).

Perspective taking. This 7-item Perspective taking subscale of the Interpersonal Reactivity Index (Davis, 1983) measures the tendency to adopt the point of view of other people in everyday life. Two versions of this subscale were used in this study. First, all participants completed the original version that measured general perspective taking ability (across targets). Second, dating partners completed a modified version that assessed their specific ability to take the perspective of their partner concerning their two nominated problems in their romantic relationship. Both versions attained reasonable internal reliability coefficients of .74 in this study.

Prior self-disclosure between the friend and the dating partners. The friend rated two statements for each problem concerning the frequency of previous problem-specific discussions between the friend and the dating partners: How often have you talked with the male dating partner about this problem in their romantic relationship?: How often have you talked with the female dating partner about this problem in their romantic relationship? Seven point Likert scales ranging from “not at all” to “very often” accompanied these items. Square root transformations were computed on each item because of their skewed distributions. Disclosure between the male partner and the friend, and the female partner and the friend, was correlated across the two problems .56 and .70 respectively. Hence, two summed variables (across the two partners) were

created to represent prior disclosure between the friend and the male dating partner, and between the friend and the female dating partner.

Verbal intelligence. Participants were administered the 29-item linguistic version of the Advanced BL-BQ, an intelligence test produced by the Australian Council for Educational Research (1982). The test measures verbal intelligence by assessing participants' vocabulary, comprehension, and verbal reasoning skills. There is a strong speed component because the time limit of 15 minutes is rather tight. It has demonstrated excellent reliability and good validity in previous research (A.C.E.R., 1982).

Personality measure. The 20-item bipolar scale used was derived from a 50-item scale developed by Goldberg (1992) to measure the Big-Five dimensions of personality (Norman, 1963). Four items were selected from each one of the Big-Five personality factors, based on the factor loadings obtained in the factor analysis reported by Goldberg (1992). All items used 9-point bipolar scales, ranging from (1) *very* (Trait A), through (5) *neither* (Trait A) *nor* (Trait B), to (9) *very* (Trait B). Examples of items measuring each factor were: silent-talkative (Extraversion); uncooperative-cooperative (Agreeableness); disorganized-organized (Conscientiousness); insecure-secure (Emotional Stability); and unintelligent-intelligent (Intellect). The instructions stressed the accuracy, honesty, and confidentiality of participant's responses.

Each dating partner filled out three copies of the scale. One scale referred to the partner's (perceiver's) own personality, another scale pertained to his or her partner's (target's) personality, and the final scale related to the perceiver's stereotypical personality rating of a member of the opposite sex of a similar age to his or her target. Each friend and stranger completed five copies of the personality scale. One scale pertained to their own personality, two further scales referred to the male and female dating partner's personality, and the last two scales related to the friend's and stranger's stereotypical personality ratings for both males and females of a similar age to the dating partners.

The results of an exploratory factor analysis (principal components analysis with orthogonal varimax rotations) of the 20 self personality scale items across the entire sample ($N = 200$) were very consistent with the Big-5 a priori factor structure.³ Five

³ As can be seen in Appendix A a few items loaded on more than one factor – a reasonably common occurrence (e.g., Goldberg, 1992).

factors were produced with eigenvalues greater than 1, together explaining 58.8% of the variance.

Results

The results of study 1 will be presented in five parts. First, I will present descriptive analyses of the major variables. Second, I will report the results of several ANOVAs that test whether empathic accuracy varies as a function of sex of the perceiver and the level of acquaintanceship. Third, correlational analyses will be presented that determine whether individual differences in perceiver's ability and target's readability exist across differing levels of acquaintanceship. Fourth, I will report the correlates of empathic accuracy at each level of acquaintanceship. Finally, I will describe the results of a series of multiple regression, mediating, and moderating analyses, that determine the predictors of empathic accuracy at each level of acquaintanceship.

Descriptive Analyses

Dating partners as perceivers. Table 1 reports the means and the standard deviations of the major variables for the dating partners as perceivers and the correlations between dating partners. On average, both male and female dating partners attained higher scores for shared cognitive focus and assumed similarity than for empathic accuracy with their respective partners (a common finding). Correlations between dating partners exhibited moderate to strong concordance for relationship satisfaction, problem-specific perspective taking, shared cognitive focus, assumed similarity, and empathic accuracy when the partner was the target.⁴ As shown in Table 2, both friends and strangers attained similar levels of empathic accuracy when judging the male and female dating partners. The same analysis was carried out to assess whether perceivers tended to achieve similar levels of empathic accuracy for the husbands and wives of the marital interaction tape. Recall that the marital interaction tape was derived from a previous study, so that these couples were strangers to all perceivers. As can be seen in Table 3,

⁴ The correlation between dating partners on shared cognitive focus is derived from two scores obtained from two separate parts of the same taped conflict discussion. Scores for both males and females, therefore, represent variables that are inherently couple-level scores. Accordingly, these two variables were combined for subsequent analyses to represent shared cognitive focus for each dating couple.

low to moderate (but significant) levels of similarity were attained by all perceiver groups.

Table 1

Means and Standard Deviations for Male and Female Dating Partners as Perceivers, and Correlations Between Dating Partners

Variables	Perceivers				<i>r</i>
	Male dating partner		Female dating partner		
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Observer-based ratings					
Target					
% EA for dating partner	47	16	54	17	.37*
% EA for stranger wives	42	13	47	16	.15
% EA for stranger husbands	41	14	49	11	.26
% Shared cognitive focus	68	18	71	18	.34*
% Assumed similarity	60	19	61	17	.46*
Behavioural diagnosticity	4.48	.89	4.39	.96	-.21
Self-report and individual difference ratings					
Attributional complexity	5.09	.94	5.55	.78	.00
Verbal intelligence	20.52	5.43	20.20	4.53	.22
General perspective taking	4.76	.98	4.74	.88	.09
Problem-specific perspective taking	5.14	.72	5.05	.98	.30*
Big-Five					
Extraversion	4.78	1.03	5.25	.92	-.13
Agreeableness	5.53	.77	5.60	.68	-.13
Conscientiousness	4.63	1.02	5.04	1.09	.22
Emotional stability	4.37	.89	3.78	.87	.05
Intellect	5.45	.54	5.48	.69	-.26
Relationship satisfaction	4.90	.69	4.31	.61	.58*
Partner closeness	5.48	1.22	5.18	1.16	.21

Note. Except for the percentage variables and Verbal intelligence, all figures were converted to a 7-pt. scale to improve readability.

EA refers to empathic accuracy.

* $p < .05$, two tailed.

Table 2
Zero-order Correlations Between Empathic Accuracy Scores for the Dating Partners as Targets and the Friends and Strangers as Perceivers

EA for male dating partner as target attained by	EA for female dating partner as target attained by	
	Friend	Stranger
Friend	.38*	---
Stranger	---	.36*

Note. EA refers to empathic accuracy.
* $p < .05$. two tailed.

Table 3
Zero-order Correlations Between Empathic Accuracy Scores Across the Male and Female Married Targets, for the Four Groups of Perceivers: Male Partners, Female Partners, Friends, and Strangers

Perceivers	Correlations across male and female married targets
Male dating partner	.31*
Female dating partner	.35*
Friend	.32*
Stranger	.54*

A problem emerges when key variables are correlated across partners, because such data violate the statistical assumption of independence. To deal with this issue that occurred primarily with the dependent variables in the current study, I aggregated the empathic accuracy scores derived for male and female targets to produce a dyad-level measure for the dating couples and the stranger married couples.⁵ In all analyses where this was done, separate analyses were also carried out on male and female targets. The results were very similar. Therefore, couple empathic accuracy scores were used for all analyses, except for when dating partners inferred each other's thoughts and feelings. In this case, by necessity empathic accuracy scores were kept at the individual level of analysis, and the issue of non-independence was statistically controlled for by the use of regression. More details will be provided when these particular analyses are presented.

Friends and strangers as perceivers. The means and standard deviations of the important variables for friends and strangers as perceivers are shown in Table 4. As can be seen, for both friends and strangers, empathic accuracy mean scores were very similar across all targets: male dating partner, female dating partner, stranger wives, and stranger husbands. Notably, in Tables 1 and 4 each of the empathic accuracy measures attained by dating partners, friends, and strangers exhibited good variability. Without such variability, the quest for individual differences in empathic accuracy would be pointless.

⁵ Relationship satisfaction was the only independent variable that was significantly correlated across dating partners. To determine whether this was a problem, I regressed female relationship satisfaction on female empathic accuracy, while controlling for male relationship satisfaction and male empathic accuracy. Similarly, male relationship satisfaction was regressed on male empathic accuracy, while controlling for both the female partner's relationship satisfaction and empathic accuracy scores. In both cases, the link between relationship satisfaction and empathic accuracy was low and non-significant; that is, the results were unchanged.

Table 4
Means and Standard Deviations for Friends and Strangers as Perceivers

Variables	Perceivers			
	Friend		Stranger	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Observer-based ratings				
<u>Target</u>				
% EA for male dating partner	42	18	38	17
% EA for female dating partner	43	17	40	21
% EA for stranger wives	44	15	40	21
% EA for stranger husbands	42	13	40	17
Self-reports and individual difference ratings				
Attributional complexity	5.42	.78	5.35	.78
Verbal intelligence	20.66	5.03	22.20	5.04
General perspective taking	4.68	.90	4.85	.96
Big-Five				
Extraversion	4.73	1.16	4.76	.86
Agreeableness	5.70	.65	5.23	.66
Conscientiousness	4.55	1.11	4.20	.99
Emotional stability	3.86	.82	3.92	.78
Intellect	5.30	.62	4.90	.64
Closeness to male partner	3.52	1.80	----	----
Closeness to female partner	4.50	1.66	----	----
Disclosure between friend & male dating partner	3.04	.90	----	----
Disclosure between friend & female dating partner	3.87	1.08	----	----

Note. Except for the percentage variables and Verbal intelligence, all figures were converted to a 7-pt. scale to improve readability.
 EA refers to empathic accuracy.

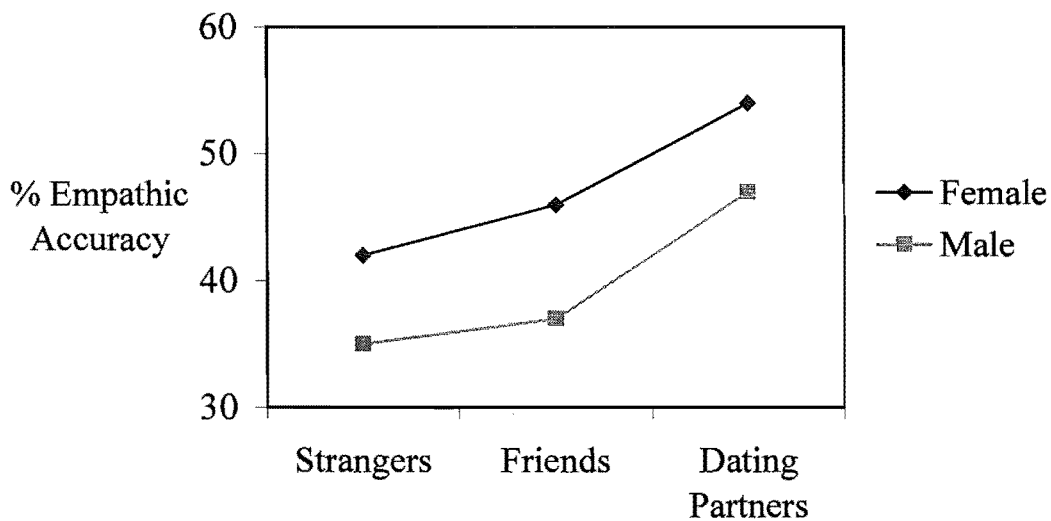


Figure 4. Mean empathic accuracy scores for partners as targets across male and female perceivers who were strangers, friends, or partners.

Empathic Accuracy: Acquaintanceship and Sex Differences

It was predicted that a) empathic accuracy would improve as a function of increased acquaintanceship, and b) women would achieve higher levels of empathic accuracy compared to men. To test these hypotheses a 3 (stranger vs. friend vs. dating partner) \times 2 (female perceiver vs. male perceiver) between subjects analysis of variance was calculated, with perceivers' empathic accuracy scores for the dating partner(s) as the dependent variable.⁶ The relevant data, depicted in Figure 4, are consistent with my predictions. The analysis revealed significant main effects for acquaintanceship $F(2, 194) = 11.66, p < .001$; and sex of perceiver $F(1, 194) = 10.21, p < .001$. A Tukey post hoc test revealed that dating partners attained significantly higher levels of empathic accuracy ($M = 50.65$) than both friends ($M = 41.18$) and strangers ($M = 38.95$). The

⁶ Couple-level empathic accuracy scores ($N=50$) were used as the dependent variable for both friends and strangers, whereas for dating partners, partner-level empathic accuracy scores ($N=100$) were utilised. Sex of the target could not be included as an independent variable in this analysis because for dating partners it is confounded with sex of the perceiver. However, a 2 (stranger vs. friend) \times (female target vs. male target) between subjects analysis of variance was calculated to determine whether empathic accuracy differed as function of the sex of the target. No main or interaction effects were found, suggesting that sex of target is not an important variable.

mean difference between friends and strangers, however, was not significant. In addition, female perceivers were generally more accurate when inferring the thoughts and feelings of the dating partner(s) ($M = 47.38$) than were male perceivers ($M = 39.81$). There was no hint of an interaction between acquaintanceship and sex of the perceiver, $F(2, 294) = .12, n.s.$

Alternative explanation. Note that the perceivers at all three levels of acquaintanceship judged the same target, and therefore differences in target's readability cannot account for the acquaintanceship effect. However, this was not a true experiment because perceivers were not randomly assigned to each group. Hence, one alternative explanation for this effect is in terms of dispositional qualities of the perceiver that could plausibly covary both with levels of acquaintanceship and empathic accuracy. For example, in the present study there was (naturally) a higher proportion of people who were currently in a romantic relationship in the dating sample, than in the friend and stranger samples. It is conceivable that participants in relationships, compared with those not in relationships, may possess better social skills or empathic ability that generalises across different targets. A strong test of this alternative explanation was to perform the same ANOVA already reported, but to control for the perceiver's empathic accuracy scores of the stranger married couples. This test should effectively control for individual differences in empathic ability that happens to be related to group membership. When this was done, the critical main effect for acquaintanceship remained significant $F(2, 193) = 9.67, p < .001$. This analysis indicates that the acquaintanceship effect is robust, and is not simply a function of individual differences in empathic ability that happens to be related to group membership.

Finally, I performed an independent t -test to test whether female's superior empathic accuracy was maintained when perceiving the stranger married couples in Session 2 of this study (using the full sample of 200 participants). Consistent with the previous analysis, female perceivers demonstrated greater empathic accuracy ($M = 46.47$) than male perceivers ($M = 39.21$), $t(198) = 3.49, p < .001$, two tailed.

To summarise, as predicted, empathic accuracy improved as a function of greater acquaintanceship. This result remained robust when differences in general empathic ability in each group of perceivers was controlled for. In addition, and in line with my prior hypothesis, female perceivers attained superior levels of empathic accuracy

compared to male perceivers when judging both the dating partner(s) and the stranger married couples.

Individual Differences in Target's Readability: The Quest for the Good Target

Recall that two target-related hypotheses were advanced in this study. The first hypothesis was that the good target across the full range of acquaintanceship (i.e. dating partners, friends, and strangers) would not be found. The second prediction was that the absence of individual differences in target's readability could be explained by Thomas and Fletcher's (1997) argument that perceiver's reliance on observational data to drive their empathic judgements should substantially decrease with greater acquaintanceship. To test the first hypothesis I computed an intraclass correlation of the three perceivers' empathic accuracy scores for both the male and the female dating partner. As expected, there was no evidence of overall consensus between the dating partner, the friend, and the stranger when perceiving the male target (intraclass $r = .23$, ns) and the female target (intraclass $r = .12$, ns).

To test the second hypothesis, perceiver's empathic accuracy was correlated with the diagnosticity of the target's behaviour. As expected, the general pattern of correlations in Table 5 showed that as acquaintanceship increased perceivers' use of behavioural information decreased. The only exception to this trend was the moderate correlation between female partner's empathic accuracy and male partner's behavioural diagnosticity.

Table 5 reveals that dating partners and friends relied on data-driven judgements to a similar extent (especially when judging the male target). Therefore, it is plausible that individual differences in target's readability occurred only when perceivers were well-acquainted with the target. To test this possibility Pearson correlations were calculated between the empathic accuracy scores attained by the male dating partner and the friend, the male dating partner and the stranger, and the friend and the stranger when the female dating partner was the target. Similarly, correlations were computed between the empathic accuracy scores achieved by the female dating partner and the friend, the female dating partner and the stranger, and the friend and the stranger when the male dating partner was the target. Table 6 shows that significant consensus between empathic accuracy scores was attained by the dating partner and the friend when

perceiving the male target (but not the female target). This result indicates that at least for well-acquainted perceivers the good target does exist.

Table 5

Zero-order Correlations Between Perceivers' Reliance on Diagnostic Behaviour and Empathic Accuracy Attained by Dating Partners, Friends, and Strangers

Perceiver	Target	
	Male dating partner	Female dating partner
Male dating partner	---	.22
Female dating partner	.41*	---
Friend	.39*	.38*
Stranger	.54*	.65*

Table 6

Target: *Zero-order Empathic Accuracy Correlations Across Different Perceiver Groups with the Male and Female Dating Partners as Targets*

Perceivers	Consensus	
	Male dating partner as target	Female dating partner as target
Dating partner-friend	.50*	.22
Dating partner-stranger	.25	.06
Friend-stranger	.25	.12

To summarise these findings, as hypothesised, individual differences in target’s readability across the full range of acquaintanceship were not found. This result can (in part) be explained by the predicted finding that perceiver’s reliance on observational data to drive their empathic judgements decreased with greater acquaintanceship. However, when empathic judgements were generated solely by well-acquainted perceivers, some evidence for the good male (but not female) target was revealed.

Individual Differences in Perceiver’s Empathic Ability: The Quest for the Good Judge

Another possible source of empathic accuracy is an underlying ability component of the judge that generalises across targets. Table 7 portrays for each class of perceiver the stability of empathic performance across targets. As predicted, the highest level of cross-target consistency occurred when perceivers judged targets whom they did not know. However, a more definitive test of perceiver’s general empathic ability is whether consistent performance is maintained across substantially different perceiver-target relationships. As seen in Table 7, moderate to substantial cross-target stability occurred regardless of the level of acquaintanceship between the perceiver and the target. Having gathered good evidence for the existence of individual differences in perceiver’s empathic ability, I next investigated the characteristics of the good judge.

Table 7
Perceiver: Correlations Between Perceiver’s Empathic Accuracy for Dating Partner(s) and Perceiver’s Empathic Accuracy for Stranger Married Couples

Perceivers	Cross-target <i>r</i>
Male dating partners	.42*
Female dating partners	.46*
Friends	.32*
Strangers	.71*

Correlates of Empathic Accuracy Attained by Each Class of Judge

Dating partners as perceiver. The correlations of all predictor variables with male and female dating partner empathic accuracy, for both their respective partners and the stranger married couples, are shown in Table 8.⁷ For male and female dating partners, higher levels of attributional complexity was associated with greater empathic accuracy when perceiving both targets (although, the correlation between male complexity and empathic accuracy for the female dating partner did not reach statistical significance). Interestingly, higher levels of problem-specific disclosure to the friend were highly correlated with superior empathic performance in female but not male dating partners. In contrast, perceptions of increased closeness to the female dating partner were significantly correlated with greater empathic accuracy attained by males, but not by females. However, empathic accuracy attained by both dating partners was unrelated to assumed similarity, verbal intelligence, general perspective taking, problem-specific perspective taking and relationship satisfaction. Finally, as expected, couples who generated higher levels of shared cognitive focus also achieved substantially higher levels of empathic accuracy when judging their partner.

⁷ None of the socio-demographic variables (e.g., education and age) were significantly correlated with empathic accuracy attained by any class of judge.

Table 8

Zero-order Correlations of Predictor Variables With Male and Female Dating Partners' Empathic Accuracy for Both Their Respective Partners and the Stranger Married Couples as the Dependent Variables

Predictor Variable	Male dating partner EA for		Female dating partner EA for	
	Female Partner	Stranger Couples	Male Partner	Stranger Couples
Male dating partner				
Assumed similarity	.00	.00	--	--
Attributional complexity	.20	.35*	--	--
Verbal intelligence	.00	.04	--	--
Disclosure to friend	.14	.02	--	--
Relationship satisfaction	.04	-.05	--	--
Partner closeness	.31*	-.03	--	--
Perspective taking				
General	.06	-.03	--	--
Problem-specific	.17	.22	--	--
Big-Five				
Extraversion	.24	.03	--	--
Agreeableness	.01	-.05	--	--
Conscientiousness	-.01	-.18	--	--
Emotional stability	-.10	-.23	--	--
Intellect	.14	-.14	--	--
Female dating partner				
Assumed similarity	--	--	.05	.07
Attributional complexity	--	--	.36*	.37*
Verbal intelligence	--	--	.10	.07
Disclosure to friend	--	--	.45*	.17
Relationship satisfaction	--	--	.26 ¹²	.15
Partner closeness	--	--	-.06	-.03
Perspective taking				
General	--	--	.21	.16
Problem-specific	--	--	.04 ¹²	-.09
Big-Five				
Extraversion	--	--	.14	.21
Agreeableness	--	--	.17	.10
Conscientiousness	--	--	.22	.26
Emotional stability	--	--	.27	.29*
Intellect	--	--	.24	.27
Couple level				
Shared cognitive focus	.45*	.02	.45*	.22

Note. EA refers to empathic accuracy. * $p < .05$, two tailed.

Friends and strangers as perceiver. Table 9 presents the correlations of predictor variables with friend and stranger empathic accuracy for both targets. For friends perceiving the dating couple, empathic accuracy was positively related to problem-specific disclosure by female, but not male, dating partners (which parallels the same finding for dating partners mentioned previously). Similarly, those friends that attained higher levels of empathic accuracy also reported higher levels of closeness to the female, but not the male, dating partner. Of the Big-Five personality factors reported by friends and strangers, only stranger's conscientiousness was positively associated with empathic accuracy for both the dating and stranger married couples. In contrast to the results previously described for dating partners in Table 8, friend and stranger attributional complexity was unrelated to empathic accuracy in either the dating or the stranger married couples.

Of particular interest is the pattern of correlations between verbal intelligence and empathic accuracy across different levels of acquaintanceship. As seen in Table 9, superior verbal intelligence was associated with greater empathic accuracy when perceivers and targets were strangers. However, when friends judged the dating couple who they knew well, verbal intelligence and empathic accuracy were unrelated.

Summary. Despite the evidence for general individual differences in perceiver's empathic ability, the correlational results show that the characteristics of the good perceiver vary markedly across (and sometimes within) different levels of acquaintanceship. For those perceivers who were well acquainted with the target, relationship-specific variables such as shared cognitive focus, perceiver-target closeness, and the target's problem-specific disclosure to the perceiver, were important predictors of empathic accuracy (although, the exact pattern of correlations varied according to the class of the perceiver). In addition, the dispositional variable of attributional complexity predicted empathic accuracy attained by dating partners, but not friends. In contrast, when perceivers judged strangers, verbal intelligence was a consistent predictor of empathic accuracy, except when dating partners were the perceivers. Other dispositional correlates of empathic accuracy with strangers as the target included conscientiousness and attributional complexity (although, again there was some variation depending on the class of perceiver).

Table 9

Zero-order Correlations of Predictor Variables With Friends' and Strangers' Empathic Accuracy for Both the Dating Couple and the Stranger Married Couples as the Dependent Variables

Predictor variable	EA for	
	Dating Couple	Stranger Couples
Friend		
Attributional complexity	.14	-.04
Verbal intelligence	.07	.33*
Disclosure between friend & female dating partner	.53*	.19
Disclosure between friend & male dating partner	-.03	.08
Closeness to female dating partner	.44*	-.15
Closeness to male dating partner	-.02	.07
General perspective taking	-.06	.13
Big-Five		
Extraversion	.02	-.14
Agreeableness	.00	-.18
Conscientiousness	-.02	-.07
Emotional stability	-.19	-.03
Intellect	.12	-.05
Stranger		
Attributional complexity	-.07	.06
Verbal intelligence	.39*	.29*
General perspective taking		
Big-Five		
Extraversion	.02	.24
Agreeableness	-.09	.00
Conscientiousness	.28*	.33*
Emotional stability	-.13	-.10
Intellect	.01	-.03

Note. EA refers to empathic accuracy.

* $p < .05$, two tailed.

Multiple Regression Analyses of Empathic Accuracy attained by Each Class of Judge

One problem with interpreting the previous correlational results is the possibility that some of the independent measures share variance. To control for shared variance, multiple regression analyses were calculated when two or more predictor variables were correlated with empathic accuracy scores for a given target. Variables that attained significant zero-order correlations were used as the predictor variables.

Predictors of empathic accuracy attained by dating partners. Table 10 presents the regression coefficients derived from a hierarchical multiple regression analysis with the male dating partner's empathic accuracy scores, when judging his partner, as the dependent variable. The male's rating of closeness to his dating partner was entered in the first set, and shared cognitive focus was entered in the second set. This order was established on the basis of a plausible *a priori* causal sequence (Cohen & Cohen, 1983). To statistically control for the influence of non-independence between the male and the female partner empathic accuracy scores, the female's score was also entered in Set 1. Consistent with the previous correlational analysis, higher levels of both closeness to the female dating partner and shared cognitive focus were associated with higher levels of empathic accuracy attained by the male dating partner.

The significant zero-order predictors of the female's empathic accuracy for her dating partner were entered into a hierarchical multiple regression (see Table 11). In accordance with the postulated *a priori* causal sequence, female attributional complexity was entered in Set 1, followed by the female's problem-specific disclosure to the friend in the second set, with shared cognitive focus inserted in the final set. To control for the effect of empathic accuracy attained by the male dating partner during the interaction, this variable was also entered in Set 1. The results were entirely consistent with the prior zero-order analysis - female empathic accuracy for her partner was positively related to female attributional complexity, her disclosure to the friend, and shared cognitive focus.

Table 10
Regression Coefficients Derived From Hierarchical Multiple Regression Analysis With Male Dating Partners' Empathic Accuracy Scores for his Partner as the Dependent Variable

Independent variables	Male dating partner EA for female partner
	Set 1
Male closeness to partner	.27*
Female partner EA	.34*
	Set 2
Shared cognitive focus	.38*
Multiple R	.57*

Note. EA refers to empathic accuracy.
* $p < .05$. two tailed.

Table 11
Regression Coefficients Derived From Hierarchical Multiple Regression Analysis With Female Dating Partners' Empathic Accuracy Scores for her Partner as the Dependent Variable

Independent variables	Female dating partner EA for male partner
	Set 1
Female Attributional complexity	.34*
Male dating partner EA	.35*
	Set 2
Disclosure between friend & female dating partner	.37*
	Set 3
Shared cognitive focus	.29*
Multiple R	.65*

Note. EA refers to empathic accuracy.
* $p < .05$. two tailed.

Mediational analysis. Given the likely a priori causal sequence of the predictors in the two previous regression analyses, I next explored the following two possibilities. First, that either the amount of problem-specific disclosure by the female dating partner to the friend or the level of shared cognitive focus, mediated the link between female attributional complexity and female empathic accuracy for the male dating partner. Second, that the amount of shared cognitive focus attained during the interaction mediated the relationship between male closeness to his dating partner and male empathic accuracy for his partner.

To be fully supported, mediational models need to satisfy three conditions (see Baron & Kenny, 1986). First, the predictor variable should significantly predict the mediating variable and the dependent variable. Second, the path between the predictor variable and the dependent variable should drop to non-significant levels (and as close to zero as possible) when the mediating variable is controlled for. Finally, the path from the mediating variable to the dependent variable should remain significant, when the predictor variable is controlled for.

The notion that the level of problem-specific disclosure to the friend by the female dating partner, prior to the interaction, mediated the link between her attributional complexity and her empathic accuracy for the male dating partner was fully supported (see Figure 5).⁸ Attributional complexity significantly predicted empathic accuracy when the mediating variable was omitted, but dropped to non-significant levels when the mediating variable was controlled for. In addition, as can be seen in Figure 5, the path from attributional complexity to prior disclosure to friend, and the path from prior disclosure to friend to empathic accuracy, were also significant. In short, female dating partners with higher levels of attributional complexity disclosed higher levels of information about the relationship problems to the friend, which in turn produced greater levels of empathic accuracy when judging her dating partner.

⁸ The influence of male partner empathic accuracy was controlled for in this path analysis.

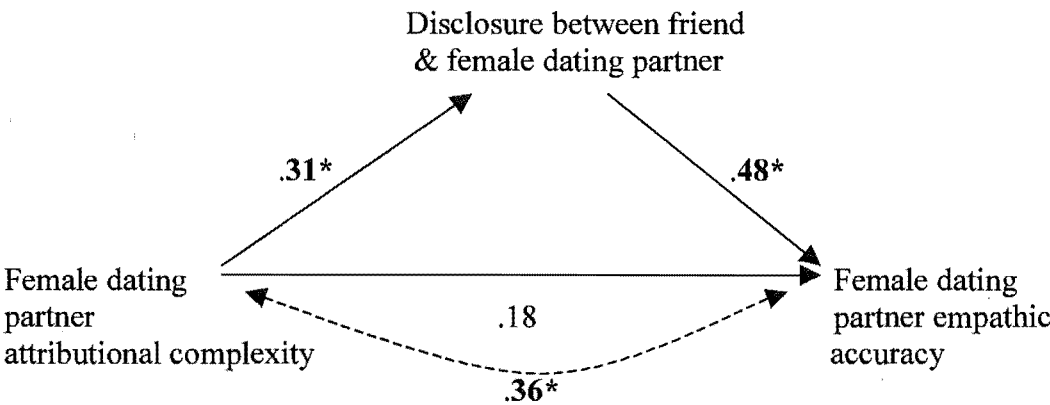


Figure 5. A path analysis depicting the relations between female dating partners’ attributional complexity and female dating partners’ empathic accuracy, as mediated by female dating partners’ problem-specific disclosure to the friend.

In contrast, the path analyses involving shared cognitive focus as the mediator for the link between either the male’s closeness to his partner or the female’s attributional complexity levels and empathic accuracy for the respective dating partner were clearly not supported. Both male’s closeness to his dating partner and female’s attributional complexity significantly predicted empathic accuracy for the dating partner when the mediating variable was omitted, but these paths remained significant when shared cognitive focus (the mediating variable) was controlled for. In addition, the paths between both male closeness to his dating partner and female attributional complexity and the mediating variable, shared cognitive focus, were not significant.

Predictors of empathic accuracy attained by friends. Again, on the basis of the plausible a priori causal order, I investigated the notion that the amount of disclosure between the female dating partner and the friend mediated the link between female partner-friend closeness and friend empathic accuracy when perceiving the dating couple. As seen in Figure 6, this mediational model was supported. Female dating partner-friend closeness significantly predicted empathic accuracy levels but dropped to nonsignificant levels when the mediating variable was controlled for. Furthermore, the path from female partner-friend disclosure to empathic accuracy for the dating couple was also significant. Therefore, friends who were closer to the female dating partner

discussed the relationship problems in greater depth with the female partner, which in turn led to higher levels of empathic accuracy attained by the friend when perceiving the dating couple.

Given that previous analyses have demonstrated that sex of the perceiver is a significant predictor of empathic accuracy, I repeated the analysis shown in Figure 6 controlling for sex. The results were virtually identical, with the significance levels remaining unchanged.

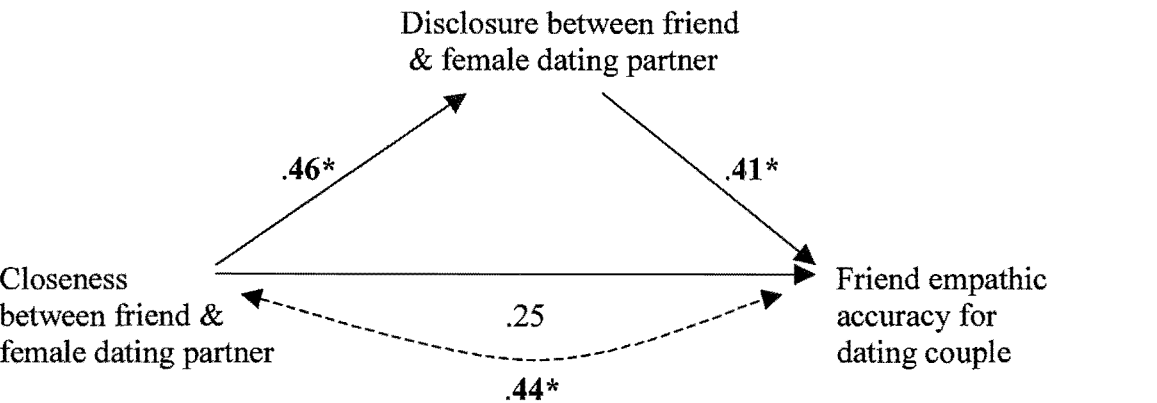


Figure 6. A path analysis depicting the relations between friends’ closeness to the female dating partner and friends’ empathic accuracy for the dating couple, as mediated by female dating partners’ problem-specific disclosure to the friend.

Predictors of empathic accuracy attained by strangers. Two simultaneous multiple regressions were calculated to determine the predictors of stranger’s empathic accuracy when perceiving the dating couple and the married couple. The relevant data are shown in Table 12. Those strangers with higher levels of verbal intelligence and conscientiousness attained higher levels of empathic accuracy when judging both target couples. However, when these regressions were repeated with sex entered as a covariate, the link between verbal intelligence and empathic accuracy remained significant, but the association between stranger’s conscientiousness and empathic accuracy levels for the dating couple ($\beta = .20$) and the married couple ($\beta = .23$) dropped to nonsignificant levels. In short, this latter relationship appears to be an artifact of sex differences in empathic accuracy.

Table 12
Regression Coefficients Derived From Two Simultaneous Multiple Regression Analyses With Strangers' Empathic Accuracy Scores for the Dating Couple and the Stranger Married Couples as the Two Dependent Variables

Independent variables	Stranger EA for dating couple	Stranger EA for married couples
Stranger		
Verbal Intelligence	.37*	.27*
Conscientiousness	.26*	.32*
Multiple R	.46*	.43*

Note. EA refers to empathic accuracy.
 * $p < .05$, two tailed.

Summary. Overall, the results of the regression analyses on the predictors of empathic accuracy attained by each group of perceivers were consistent with the previous correlational analyses. For the male dating partner, higher levels of closeness to the female dating partner and shared cognitive focus were associated with higher levels of empathic accuracy for his partner. For the female dating partner, higher levels of attributional complexity, prior problem-related disclosure to the friend, and shared cognitive focus, were related to empathic accuracy when perceiving her partner. However, further analysis revealed that the level of prior problem-specific disclosure to the friend by the female dating partner mediated the link between her attributional complexity and her empathic accuracy for the male dating partner. The important mediating role played by such disclosure between the friend and the female dating partner was also evident in predicting empathic accuracy attained by friends. That is, friends who were closer to the female dating partner discussed the relationship problems in greater depth with the female partner, which in turn led to higher levels of empathic accuracy attained by the friend when perceiving the dating couple. Finally, controlling for sex of the perceiver, strangers with higher levels of verbal intelligence achieved superior levels of empathic accuracy when perceiving both the dating and the married couples.

Exploring the Role of Attributional Complexity and Verbal Intelligence in Predicting Empathic Accuracy across Different Levels of Acquaintanceship

One interesting and unexpected trend to emerge from this study were the complementary roles of verbal intelligence and attributional complexity in predicting empathic accuracy, which systematically varied across different levels of acquaintanceship. In general, verbal intelligence was an important predictor when perceiving strangers, whereas attributional complexity was a major predictor when dating partners judged each other. Note that, replicating past research findings, attributional complexity and verbal intelligence were not correlated ($r = -.08$). This pattern of results implies that as the level of acquaintanceship increased the predictive value of attributional complexity was enhanced, whereas the predictive value of verbal intelligence declined. Admittedly there was one major exception to this trend; namely, when the dating couples judged the stranger married couples, empathic accuracy was related to attributional complexity, not verbal intelligence. A full explanation for the various findings will be presented in the discussion section. However, to give a truncated version here, a good way to conceptualise acquaintanceship, is in terms of the quantity and quality of information or knowledge available to the perceiver concerning the target (Funder, 1995). Hence, my explanation for this trend is in terms of two related hypotheses. First, that verbal intelligence is a good predictor under conditions of novelty or when limited information was available (i.e. when judging strangers). Second, that attributional complexity is a good predictor under conditions of high information (i.e. when judging well acquainted targets), where perceivers are required to effectively combine their elaborate pre-existing theories of the target with the incoming observational data evinced during the interaction.

As already noted, I found that neither attributional complexity nor verbal intelligence was correlated with friend empathic accuracy when perceiving the dating couple. The previously stated explanation suggests that these null results could be explained by the possible moderating influence of the amount or quality of information available to the friends concerning the target. Friends in this study varied on a good indicator of information quality - the amount of information dating partners disclosed to the friend regarding the relationship problems. Hence, one possible explanation for the

null finding between attributional complexity (or verbal intelligence) and friend empathic accuracy, was the moderating influence of the amount of dating couple-friend disclosure. To be consistent with the other findings, the analysis should reveal a positive relation between attributional complexity and empathic accuracy at high levels of information, and a negative or null relation at low levels of knowledge.

To test this explanation, the measures of disclosure to the friend by the male, and the female dating partner, were aggregated to represent a combined score for problem-specific disclosure to the friend. Median splits were used to form High and Low information groups for disclosure. Next, a multiple regression equation was calculated to test the moderating influence of levels of knowledge (see Baron & Kenny, 1986). Attributional complexity, dating couple-friend disclosure, and the interaction term (attributional complexity \times dating couple-friend disclosure) were entered simultaneously into the equation. The analysis revealed a significant contribution by the interaction term to the prediction of friend empathic accuracy for the dating couple ($t = 2.77, p < .01$)⁹. To illustrate the nature of this interaction, the standardised regression slopes for attributional complexity as a function of high and low disclosure are depicted in Figure 7. As expected, under conditions of high disclosure attributional complexity was positively related to friend empathic accuracy ($\beta = .40$), whereas under conditions of low disclosure attributional complexity was negatively associated with friend empathic accuracy ($\beta = -.27$). This result provides good evidence for the idea that as the level of acquaintanceship increases, higher levels of attributional complexity are associated with superior empathic accuracy.

⁹ This is a particularly compelling result given the fact that reliable moderator effects are notoriously hard to attain (Paunonen, 1989).

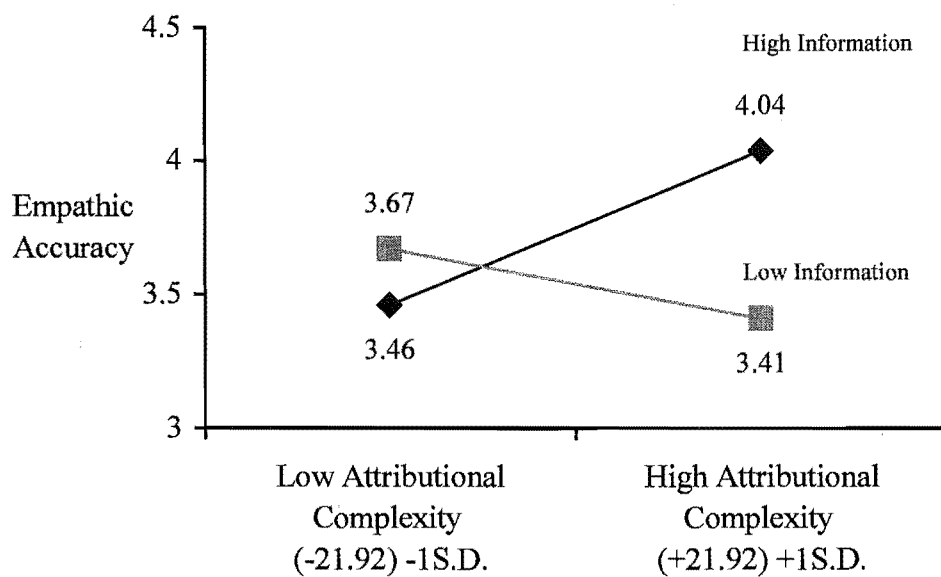


Figure 7. Friend's empathic accuracy for the dating couple as a function of disclosure to the friend by the female dating partner and attributional complexity.

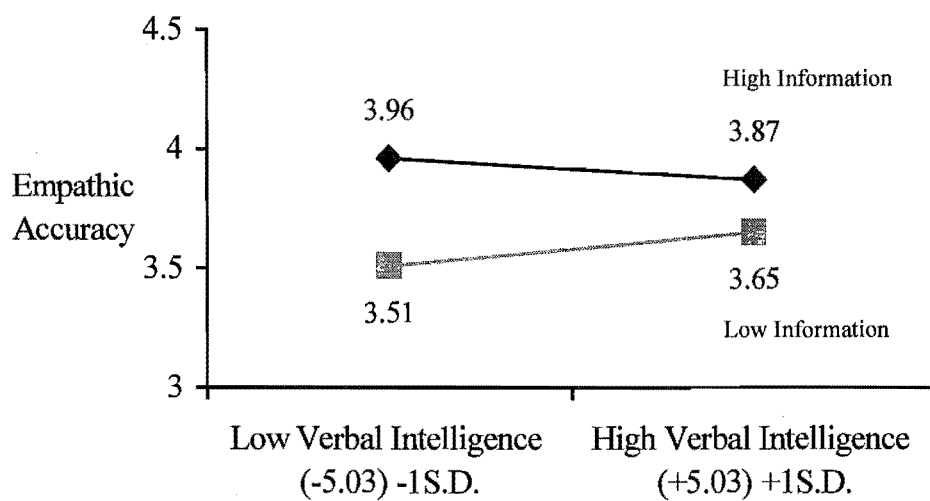


Figure 8. Friend's empathic accuracy for the dating couple as a function of disclosure to the friend by the female dating partner and verbal intelligence.

The next question considered was to what extent the relationship between *verbal intelligence* and friend empathic accuracy was also moderated by the measure of disclosure. As previously, the regression analysis was conducted, but with verbal intelligence replacing attributional complexity as a main effect and as part of the interaction term in the equation. The interaction term did not contribute significantly to the prediction of friend empathic accuracy. Nevertheless, Figure 8 shows that the direction of the regression slopes were opposite to those in Figure 7, as I had predicted. Although the effects were weak, higher intelligence was associated with greater empathic accuracy, only under conditions of *low* information.

To summarise, an unexpected trend to emerge from this study was that as the level of acquaintanceship increased the predictive value of attributional complexity was enhanced, whereas the predictive value of verbal intelligence declined. My post hoc explanation for this pattern of results is in terms of the quality and quantity of information available to the perceiver concerning the target. Verbal intelligence is a good predictor of empathic accuracy under conditions of limited information or novelty, whereas attributional complexity is a good predictor under conditions of high information - where perceivers have to effectively combine pre-existent rich amounts of individuating information with observational data. Regression analyses were largely consistent with this explanation. When dating partners judged each other, higher levels of attributional complexity, rather than verbal intelligence, was associated with greater empathic accuracy (although there was some variation depending on sex). In contrast, when strangers judged the dating couples higher levels of verbal intelligence (not attributional complexity) were related to superior empathic accuracy. Furthermore, the relationship between attributional complexity and empathic accuracy attained by the friends when perceiving the dating couples was moderated by the amount of prior disclosure between the friends and the dating couples concerning the nominated relationship problems. As expected, under conditions of high disclosure attributional complexity was positively related to friend empathic accuracy, whereas under conditions of low disclosure attributional complexity was negatively associated with friend empathic accuracy. Although the test of the moderating influence of disclosure on the relationship between verbal intelligence and friend empathic accuracy did not reach significance, the direction of the regression slopes were nevertheless consistent with the

notion that verbal intelligence is associated with increased empathic accuracy under conditions of limited information.

Discussion

Summary of Findings

Relationship-based predictions. As predicted, empathic accuracy significantly improved as a function of increased acquaintanceship. Moreover, the acquaintanceship effect remained robust when individual differences in judges' empathic ability and targets' readability were eliminated or controlled for. In terms of the relationship-level predictors of empathic accuracy, the results were generally consistent with my expectations. First, dating partners who cognitively focused on the problems to a greater extent during their interaction attained higher levels of empathic accuracy. Second, higher levels of perceived relationship closeness were associated with more accurate empathic judgement (for male dating partners). Third, the beneficial effects of problem-related disclosure on empathic accuracy were demonstrated for both the discloser and the confidant. Further path analyses revealed that such disclosure mediated the link between other predictor variables and empathic accuracy in theoretically feasible ways. From the disclosers' (or female dating partners') perspective, those with higher levels of attributional complexity disclosed higher levels of information about the relationship problems to their friend, which in turn produced greater levels of empathic accuracy when the discloser judged her dating partner. Similarly, from the confidants' (or friends') perspective, those who were closer to the female dating partner were confided in to a greater extent by the female dating partner concerning the relationship problems, which in turn led to higher levels of empathic accuracy when the confidant judged the dating couple

Target-based predictions. The influence of target's readability was found to vary according to the level of acquaintanceship (dating partners, friends, or strangers). As expected, perceivers' reliance on diagnostic behaviour decreased as a function of acquaintanceship. My explanation is that as acquaintanceship increases, the empathic judgement process becomes less data-driven and more theory-driven. However, more limited evidence for the good target did emerge at moderate to high levels of

acquaintanceship, with some male targets being judged more accurately by their dating partners and friends than others.

Judge-based predictions. Compelling evidence was found for the good judge of others' thoughts and feelings in the context of a difficult empathic accuracy task. Moreover, individual differences in judges' empathic accuracy generalised across widely different levels of acquaintanceship. Furthermore, as predicted, female perceivers attained superior levels of empathic accuracy compared to male perceivers, regardless of their relationship with the target. In contrast, the association between empathic accuracy and the dispositional variables, attributional complexity and verbal intelligence varied substantially according to the nature of the judge-target relationship. Further analyses generally supported the explanation that more intelligent judges generated more accurate empathic inferences under conditions of low information (or acquaintanceship), whereas more attributionally complex judges achieved superior levels of empathic accuracy under conditions of high information (including possessing substantive prior theories about the target).

Explaining Relationship-level Effects on Empathic Accuracy

The acquaintanceship effect. As hypothesised, the basic finding showed that increased acquaintanceship led to greater empathic accuracy. This result concurs with previous research concerning the beneficial effects of greater information or knowledge on the validity of empathic judgements (e.g., Graham, 1994; Marangoni et al., 1995; Stinson & Ickes, 1992, but c/f. Hancock & Ickes, 1996), and is consistent with RAM's (Funder, 1995) theoretical postulate that the availability of "good information" is a basic moderator of judgement accuracy.

According to Thomas and Fletcher (1997), the effect of acquaintanceship on empathic accuracy can be explained in terms of the differential reliance on theory-driven vs. data-driven judgements by insiders and outsiders. Insiders (dating partners), subjective outsiders (friends), and objective outsiders (strangers) all have access to the behavioural information evinced by the targets during the interaction. However, the results of this study show that perceivers' reliance on behavioural data to drive their judgements decreases as their level of acquaintanceship with the target increases. Instead, empathic judgements made by insiders, and to a certain extent subjective

outsiders, seem to be substantially influenced by their rich and elaborate theories concerning the target and the relationship forged over a history of prior interaction.

There are at least two plausible reasons why judgements that are derived from elaborate target/relationship-specific theories attain more accuracy than judgements entirely based on the face-value behavioural information. First, insider knowledge and theory may be required to accurately interpret idiosyncratic target behaviour. For example, behaviours that mean nothing to an objective observer, such as a raised eyebrow or a vein throbbing in the forehead, may be pregnant with meaning to the partner or friend because they are interpreted in light of pre-existing target-specific lay theories and knowledge that are inaccessible to a stranger. Second, a critical feature of close relationships is the extent to which people provide diagnostic feedback and disclosure to each other concerning their reactions to the situations and events that they mutually experience (Altman & Taylor, 1973; Reis & Patrick, 1996). As such, intimates are likely to construct shared theories or knowledge structures about the meaning of various kinds of interactive behaviour (Colvin et al., 1997; Hancock & Ickes, 1996). Thus, even in the absence of diagnostic behaviour, indicating anger, for example, a man may attribute anger to his dating partner based on his detailed knowledge of her personality, or more specifically knowledge of her beliefs and feelings concerning the problem under discussion.

Upon closer examination, the results revealed that dating partners were more accurate judges than friends and strangers, but the difference between the empathic performance of friends and strangers was relatively slight. The superior performance of dating partners compared to friends is perhaps not surprising given that these judgements are made in the context of a problem-solving discussion concerning their own dating relationship. Insiders are likely to have more detailed and valid theories of each other's characteristic thoughts and feelings concerning specific relationship problems (via more extensive observation and feedback) than subjective outsiders. However, this finding is contrary to a commonly expressed proposition in the close relationship arena that individuals can be the most knowledgeable, yet the least objective judge of their partner (Sillars & Scott, 1983).

The failure of friends to significantly outperform strangers was unexpected. It was thought that subjective outsiders' reliance, in part, on their detailed pre-existing

knowledge of the target, and to a lesser extent, the dating relationship, would give them a unique advantage over the substantially data-driven judgements of objective outsiders in the context of a rich and idiosyncratic interaction. However, this result is consistent with the study by Hancock and Ickes (1996) which demonstrated no difference between the empathic accuracy attained by subjective (friends) and objective (strangers) outsiders when observing the target in the context of a highly scripted, getting-acquainted interaction with a complete stranger. In fact, previous research has only documented an acquaintanceship effect (friends vs. strangers), when friends have been the insiders, and have had the opportunity to use their insider knowledge when judging each other's thoughts and feelings during a friendship-related discussion (e.g., Graham, 1994; Stinson & Ickes, 1992). This boundary condition on the acquaintanceship effect is also consistent with the positive relationship found in the current study between friends' empathic accuracy and both the levels of closeness and disclosure reported between the friends and the female dating partners. That is, friends who were more like the dating partner (in terms of the depth and quality of their target-relevant information and theories) performed at accuracy levels closer to those attained by the dating partner than the stranger. Taken together, these results suggest that for subjective outsiders to outperform objective outsiders the judge's pre-existing target-specific knowledge structures need to be specifically relevant to the content of the interaction.

Alternative explanations. It is important to note that the research design and analyses ruled out several plausible competing explanations for the acquaintanceship effect. Dating partners did not achieve more accurate empathic judgements than friends and strangers because they (a) judged a target who was easier to read, or (b) possessed higher levels of ability in making accurate empathic judgements. However, another possible explanation, that could not be obviated in the current study, is in terms of the observational status (passive vs. active) of the judge. In this study, dating partners interacted with the target, whereas friends and strangers passively observed the target on videotape. It is conceivable that active observation may produce more accurate judgements than passive observation because the interactants can direct the discussion towards issues that they are more knowledgeable about. Two recent studies, however, have failed to find a link between the observational status of the perceiver and judgement

validity in either the trait accuracy (Blackman & Funder, 1995) or the empathic accuracy domains (Hancock & Ickes, 1996).

The characteristics of the good relationship. A general contention of this thesis was that the unique relationship between the judge and the target plays a major role in determining the accuracy of empathic judgements. Indeed this focus was justified, with relationship-level variables being the best predictors of empathic accuracy attained by well-acquainted judges.

Replicating Thomas et al. (1997), I found that shared cognitive focus was substantially positively related to increased empathic accuracy. One way to account for this finding is in terms of the previously outlined distinction between data-driven and theory-driven accuracy. Partners who put more cognitive effort into their problem-solving discussions pay closer attention to the interpersonal flow of information, and who adopt a shared frame of reference, should be more accurate because: (a) they are more able to detect and correctly utilise the straightforward diagnostic behavioural data, (b) they are more capable of perceiving and utilising the subtle idiosyncratic diagnostic cues evinced by their partner that are available only to an insider, and (c) they are more able to effectively access and utilise their pre-existing knowledge and theories concerning their partner.

An intriguing finding, given the mixed pattern of previous research results, was that relationship satisfaction was not associated with empathic accuracy. However, caution should be exercised in generalising this result to other relationship contexts. It is quite conceivable that the link between relationship satisfaction and empathic accuracy will range from negative to positive, depending on the context. For example, when the relationship is under severe threat (e.g., because of one partner's sexual infidelity), happy partners, when compared to unhappy partners, may be motivated to produce inaccurate empathic judgements (see Simpson et al., 1995). Under other conditions – for example, when discussing something positive and non-conflictual – love's rose coloured glasses may conceivably confer an advantage in terms of empathic accuracy.

In contrast to relationship satisfaction, friends (confidants) who were closer to the female dating partner discussed the problems in greater depth (with the female partner) and subsequently improved their performance at reading both the male and female dating partners. Increased levels of problem-solving discussion also improved the empathic

accuracy of the female dating partners (disclosers) for their male partners. In short, the benefits of such intimate disclosure with the friends were evident for the female dating partners and the friends, but eluded the male dating partners completely. This suggests that the kind of self-disclosure evinced by men and women about relationships is consistent with the stereotypes provided in TV shows and films. Women provide their friends with detailed, raw, uncensored, and, thus, highly diagnostic information about feelings and cognitions. In contrast, male talk about relationship problems is superficial, descriptive, and uninformative about deeper feelings, cognitions, and relationship interactions.

These findings were largely expected and are remarkably consistent with the general argument, previously expressed, that women are much more focussed on, and attentive to, the psychological dynamics of close relationships, whereas men are more individualistic and have a less sophisticated approach to the psychological workings of intimacy. Nevertheless, the results also break new ground in two respects. First, they show that the benefits of self-disclosure are revealed in direct measures of accuracy as well as self-reported measures of understanding. Second, that intimate disclosure (by females) appears to be the single most important predictor of empathic accuracy for both the discloser and the confidant.

Males who felt closer to their female dating partner attained more accurate judgements of their partners' thoughts and feelings, whereas this was not true for women.¹⁰ One speculative explanation for this unexpected gender difference, consistent with the results of the previously described path analysis, is in terms of the more extensive and diagnostic disclosure provided by females than males in intimate relationships. Males in closer relationships are likely to have greater access to detailed and diagnostic disclosure from their female dating partners about her innermost thoughts, feelings and attitudes concerning the dating relationship (including the relationship problems). Given the finding that males, in particular, seem to heavily rely on their prior theories and knowledge, this increased flow of veridical information should improve the accuracy of males' judgements. In contrast, females seem to receive the same quality and amount of information and feedback from their male partners, regardless of how close

¹⁰ It is important to note that perceived relationship closeness was not was not correlated across dating partners.

their relationship is. Hence, the accuracy of female dating partners' empathic judgements is not altered by how close they feel to their romantic partner.

Explaining Target Effects on Empathic Accuracy

As expected, individual differences in target's readability were not found across different levels of acquaintanceship, as evidenced by a lack of consensus or covariation in the empathic accuracy scores of the three classes of perceivers when judging the same targets. Recall that according to RAM (Funder, 1995), two conditions are required for target effects to emerge. First, targets must vary in the extent to which they produce available behavioural cues that are relevant to their private thoughts and feelings. In the current study this condition was satisfied, with considerable variability being exhibited in the diagnosticity of targets' behaviour. Second, judges at each level of acquaintanceship must detect and then utilise these behavioural cues when deriving their empathic judgements. However, as I have already described judges at different levels of acquaintanceship utilise behavioural cues in very different ways. Hence the lack of target effects in this study is predictable. Target effects were simply overwhelmed by the influence of the judge-target relationship in the context of rich and idiosyncratic problem solving discussions (but c/f Hancock & Ickes, 1996; Marangoni et al., 1995).

However, when attention was restricted to empathic judgements generated solely by well-acquainted perceivers, some evidence for the good target was revealed. Specifically, some male (but not female) targets were more accurately judged by their dating partner and friend than were others. This result is generally consistent with research in the trait accuracy domain that has demonstrated individual differences in target judgeability only within the confines of relatively long-standing friendships (e.g., Colvin, 1993a, 1993b).

The question arises why target effects occurred for men but not women as targets. First, female dating partners and friends relied on data-driven judgements to a similar extent when inferring the male target's thoughts and feelings. Second, as reported previously, the amount of prior problem-related disclosure by the female dating partner to the friend was an important determinant of empathic accuracy attained by both the discloser and the confidant. Such diagnostic disclosure enables female dating partners and friends to construct shared knowledge structures regarding the relationship and its

problems, which in turn may generate a higher level of consensus between the empathic accuracy attained by these well-acquainted judges. In other words, reliance on mutually held theory-driven judgements by female partners and friends may account for the male target effect. To test this explanation, I correlated female dating partner's empathic accuracy with friend's empathic accuracy (when judging the male target), while controlling for the effect of the amount of prior problem-related discussion between the female partner and the friend. The level of consensus dropped considerably (from $r = .50$ to $r = .29$), although the partial correlation coefficient remained significant. Thus, the existence of such mutually held problem-specific theories accounts for an important component of the target effect.

Explaining Judge Effects on Empathic Accuracy

The good judge. The present study provided a rigorous test of individual differences in empathic ability. The results provide support for the previously outlined argument that a difficult empathic accuracy task is a necessary condition for individual differences in judgement ability to emerge. In both the current research and the Marangoni et al. (1995) study, in which strangers generated empathic judgements in the context of behaviourally intense and complex interactions, strong evidence for the good judge has been found. In contrast, when perceivers have had the easier task of inferring targets' thoughts and feelings during highly scripted and desultory "getting acquainted" interactions between two complete strangers, judge effects have failed to emerge (see Hancock & Ickes, 1996).

Moreover, this set of findings extends previous research by demonstrating individual differences in empathic ability across different judge-target relationships. Specifically, the good judge of strangers was also shown to be a good judge of their friends and their dating partner. Thus, this study provides the first clear demonstration, in the arena of empathic accuracy, of what could be regarded as individual differences in social intelligence in the arena of empathic accuracy. Of course, this result begs the oldest question in accuracy research - what are the attributes of the good judge?

The characteristics of the good judge. As expected, judges' self-reported levels of general perspective-taking ability failed to predict empathic accuracy. This pattern of results is by no means unique. Although judges differ reliably in their ability to

accurately infer the thoughts and feelings of others, they appear to have little insight regarding their own relative levels of accuracy (see Davis & Kraus, 1997; and Ickes, 1993, for a review). Moreover, consistent with prior meta-analytic results, none of the Big-5 factors of personality were meaningfully related to empathic ability (Davis & Kraus, 1997). Although more conscientious judges were more accurate perceivers of strangers' thoughts and feelings, further analyses revealed that this relationship was a by-product of sex differences in empathic ability.

A striking feature to emerge in this study was the clear pattern of gender differences in empathic ability. Female judges at each level of acquaintanceship attained higher levels of empathic accuracy than did male judges. Moreover, this effect remained robust when the gender of the target was controlled for. This result accords well with the general theoretical viewpoint that women are inclined to forge and sustain a sense of relatedness and interdependence, whereas men seek to construct and maintain a sense of individuality and independence (Bakan, 1966; Cross & Madson, 1997). Indeed, there is good evidence that these differences are especially salient in the context of intimate romantic relationships (see Acitelli & Young, 1996; for a review). However, this finding is contrary to Thomas et al (1997) who found no hint of gender differences in the empathic accuracy attained by spouses in long-standing marriages (mean length > 15 years). One explanation for these discrepant results is that gender differences in empathic accuracy may dissipate over the course of long-standing marriages. For example, the typical husbands relatively individualistic and less sophisticated approach to marriage may be open to attack and subsequent change after long-term exposure to their wives' style of interpersonal thinking. In contrast, sex differences in the earlier stages of the relationship life cycle (which make up the bulk of the sample in this study) may well be more robust.

Despite the evidence for general individual differences in perceiver's empathic ability, the results showed that the characteristics of the good judge varied markedly according to the nature of the perceiver-target relationship. Higher levels of verbal intelligence were associated with greater empathic accuracy when perceiving strangers. In contrast, higher levels of attributional complexity were related to superior empathic accuracy when judging well-acquainted targets (although this association was not statistically significant for male dating partners). This finding is consistent with what IQ

is often thought to measure – the ability to solve novel problems. In contrast, attributional complexity is a good predictor under conditions of high information where judges have access to rich and detailed pre-existing theories of the target. As my path analysis suggests, under such conditions, more complex judges develop more extensive and accurate theories regarding the relationship and its problems, which leads to more accurate judgements. These results are also nicely consistent with previous findings that the benefits of more complex attributional schemata emerge most clearly during especially complex and information-rich judgement tasks (see Fletcher, et al., 1990; Fletcher, Grigg, & Bull, (1988).

One result that does not quite fit my interpretation was that when the dating couples judged the stranger married couples, empathic accuracy was related to attributional complexity, not verbal intelligence. A speculative explanation for this finding is in terms of the observational status of dating partners compared to that of friends and strangers. Recall that dating partners participated in their own problem-solving discussion and then subsequently observed the stranger married couples participate in a very similar interaction, whereas friends and strangers merely observed the interactions of both the dating and married couples. Perhaps the act of prior participation equipped dating partners with more detailed knowledge about the nature of the interaction and how the interactants would characteristically be thinking and feeling based on their own experiences.

Taken together, and interpreted carefully, my results are consistent with prior research and theorising but also break new ground. The current study is the first to unearth individual difference markers of empathic ability. Although previous researchers in the trait accuracy arena have cautiously implicated intelligence and cognitive complexity as characteristics of the good judge (e.g., Taft, 1955; Davis & Kraus, 1997), few studies have yielded effect sizes of the magnitude found in this study. Overall, this set of findings confirm the speculations of early accuracy researchers that the attributes of the good judge of strangers might be independent of the characteristics associated with the good judge of well-acquainted targets (Allport, 1937; Vernon, 1933). Moreover, the finding that the characteristics of the good judge vary according to the nature of the judge-target relationship probably constitute a major reason why inconsistent and weak

results related to individual differences have typically been found in the long history of accuracy research.

STUDY 2

THE MODERATORS OF TRAIT ACCURACY AND CONSENSUS

The focus of study 2 is on the accuracy and consensus of personality judgements, rather than on empathic judgements. Recall that in Study 1 in order to explain the cognitive processes used by insiders and outsiders when making empathic judgements in close intimate relationships, it was argued that RAM (Funder, 1995) needed to be augmented with a social cognitive perspective. However, the role of social cognitive processes is given a less pivotal role in the current study. RAM provides a systematic classification of moderator variables that might account for trait accuracy (and consensus). Funder (1995) proposes that the characteristics of the judge, target, trait, and the quantity and quality of available information may all affect the accuracy of trait judgements, and that interactions among these moderators are all possible. In the current study, I examined the effect of two of these basic moderators on accuracy and consensus in personality judgements – “good information” and the “good trait”, along with their unique interaction. The design of this study effectively precluded a thorough examination of the person-based moderators of trait accuracy. Individual differences in judge’s ability (the “good judge”) could not be systematically assessed because some judges (the dating partners) only perceived one target. Finally, although it was possible to collect some data on the “good target”, each target was only judged by one perceiver at each level of acquaintanceship. Hence, I was unable to cleanly separate target effects from acquaintanceship effects. In short, the design of this study does not allow as comprehensive an analysis as Study 1. However, it does allow me to test whether past findings in the trait accuracy literature extend to the context of romantic relationships, as well as test whether my prior results replicate in this study.

This introduction consists of three main sections. The first section investigated the moderating effect of good information (or the level of acquaintanceship between the judge and target) on trait accuracy and consensus. The second section focused on the impact of trait-based moderators on accuracy and consensus, that is, whether some trait domains were more easily judged than others. The final section examined whether trait

observability and acquaintanceship interacted in their effects on trait accuracy and consensus.

As mentioned previously, the choice of criterion measures for accuracy is difficult. In line with most prior studies, trait accuracy was measured in terms of self-other agreement. In other words, a judge was considered accurate to the extent that his or her description of a target person's personality characteristics matched the target's self-description. Although the practice of using self-other agreement as a criterion for trait accuracy has attracted criticism (e.g., John & Robins, 1994; Kenny, 1994), alternative criteria such as behavioural prediction are also problematic (Blackman & Funder, 1998). This issue will be revisited in more detail in the discussion.

In contrast to the measurement of trait accuracy, I departed from the conventional method of measuring consensus. Researchers have typically assessed consensus between judges at the same level of acquaintanceship (e.g., the extent to which two friends agree with each other when judging the personality of a third target person (their respective friend) or the amount of consensus attained by two strangers when judging the personality of a target person who they have just briefly met). However, in this study because only one judge was available at a given level of acquaintanceship, consensus was assessed between judges at different levels of acquaintanceship. For example, I measured the degree of consensus between well-acquainted judges (dating partner and friend) when describing the target's personality, and consensus attained when one judge was well-acquainted and the other a stranger.

Good Information: The Influence of Acquaintanceship on Trait Accuracy

Previous Theorising

Psychologists have frequently demonstrated a penchant for debunking intuitive notions of social reality. A good example of such scepticism concerns the link between acquaintanceship and accuracy. Although, common sense dictates that it takes considerable time to really get to know someone, this issue has generated extensive debate, particularly in the trait accuracy literature.

On the one hand, from an ecological vantagepoint, judgement accuracy is largely a function of whether sufficient diagnostic information is available in the *immediate*

social environment (McArthur & Baron, 1983). Thus, if accuracy is primarily data-driven, then the effect of the perceiver-target relationship on accuracy should be trivial. Similarly, Kruglanski's (1989) lay epistemic theory asserts that greater amounts of information may not improve accuracy, and occasionally may even lead judges astray. Well-acquainted judges have access to so much target-relevant information that they may sometimes engage in a form of interpretative overkill, to the detriment of judgement validity (Cloyd, 1977; Taft, 1955; Wilson & Schooler, 1991). Motivational factors have also been posited as causing inaccuracy of judgements in close relationships. Intimate judges of close others, are the most knowledgeable but also may be the least objective perceivers (Sillars & Scott, 1983). The desire to maintain a rose-tinted or charitable view of their partner or friend can presumably overwhelm the inclination to dispassionately use highly diagnostic information that is available (e.g., Murray & Holmes, 1993).

On the other hand, more recent theorising has depicted the amount of quality information as an important determinant of the accuracy of personality judgements. According to RAM, as a perceiver observes and interacts with a target on more occasions, more relevant behavioural cues become available, and hence the likelihood of the perceiver detecting and using such cues increases (Blackman & Funder, 1998; Funder, 1995). Therefore, RAM's general prediction is that as the quantity and quality of information increases the accuracy of personality judgements improves. In a similar vein, the Weighted Average Model (WAM, Kenny, 1991, 1994) predicts that greater acquaintance leads to greater accuracy (typically operationalised as self-other agreement), providing there exists at least some degree of cross-situational consistency in the target's behaviour.

Both RAM and WAM also generally predict that interjudge consensus increases with acquaintanceship, albeit for different reasons. RAM accounts for the effect of acquaintanceship on accuracy and consensus in the same way – increased acquaintance leads to access to more diagnostic or “good information” upon which the personality judgements are based. WAM's explanation is not in terms of the quantity or quality of information per se, but rather that the amount of information the two judges explicitly share or have in common with each other will influence the degree to which they reach consensus in their judgements. In general, Kenny's model predicts that higher levels of shared information leads to greater consensus. WAM identifies at least three ways that

judges can gain access to shared information: overlap, communication, and similar meaning systems.

The overlap explanation states that consensus is determined by the extent to which judges observe the same target behaviours. As acquaintanceship develops people (judges) spend more time with each other, and are also more likely to be together when they encounter mutual acquaintances (targets). During such encounters judges observe the same target behaviours and thus the information available to them considerably overlaps (Kenny & Kashy, 1994). In other words, consensus improves as a function of acquaintanceship by virtue of increased levels of behavioural overlap between judges. One caveat to WAM's prediction, however, is that under conditions of extremely high or "perfect overlap" consensus will be unrelated to acquaintanceship. That is, if judges are always privy to the same behaviour then consensus will remain constant across the life cycle of the perceiver-target relationship.

The second explanation is that prior communication between judges forges the basis for interjudge consensus. As people become better acquainted they are more likely to discuss each others' personalities. WAM posits that if these discussions mutually influence judges' perceptions of the target then consensus and acquaintanceship will be positively related (Kenny & Kashy, 1994). Kenrick and Funder (1988) refer to this as the "cahoots" hypothesis, where judges agree with each other merely because they are in cahoots with each other.

The final explanation is that consensus is attained by judges interpreting and utilising the behavioural information evinced by the target in the same way. It is conceivable that well acquainted judges are more likely to adopt a shared frame of reference and explain and label others' behaviour in more similar ways than more casual acquaintances (Kenny & Kashy, 1994). Hence, the positive relationship between acquaintance and consensus may be moderated by the extent to which judges share similar meaning systems.

Previous Research

Empirical efforts to resolve this debate have yielded mixed results when the impact of information has been assessed either at the early stages of acquaintanceship or in relationships where the range of target behaviours observed by the judge are relatively

constrained. Initial attempts to manipulate the level of the judge-target relationship experimentally, by varying the length and content of vignettes describing a hypothetical target person, not surprisingly found a null relationship between acquaintanceship and accuracy (e.g., Stelmachers & McHugh, 1964). Such artificial measures fail to capture the richness and complexity of everyday person perception (Kenny, 1994). Using a more ecologically valid manipulation of target-relevant information, Blackman and Funder (1998) had judges observe between five and thirty minutes of target's videotaped behaviour evinced during unstructured dyadic interaction. They found that after thirty minutes observation, judges generated more accurate judgements of target's visible personality traits than after five minutes. However, the level of interjudge consensus attained under conditions of "perfect overlap" (Kenny, 1991, 1994) did not change over the course of acquaintanceship.

Researchers have also used longitudinal designs to determine whether accuracy and consensus improve over the course of acquaintanceship in the context of laboratory interactions, classroom settings, and occasionally residential settings (i.e. roommates). For example, Paulhus and Bruce (1992) demonstrated that the accuracy of students' personality judgements of their fellow classmates significantly improved over the first seven weeks of class, while consensus remained unchanged (see also Paulhus & Reynolds, 1995). In contrast, two more recent longitudinal studies found no link between the length of acquaintanceship and the veridicality of people's judgements of their fellow roommates' personalities (Park, Kraus, & Ryan, 1997; Swann & Gill, 1997). Furthermore, in a review of the longitudinal studies that have measured interpersonal perception in these relatively constrained contexts, Kenny, Albright, Malloy, and Kashy (1994) found no evidence that consensus is enhanced by increasing acquaintanceship (which was explained in terms of the very high levels of information overlap between judges).

However, stronger and more consistent findings have been obtained when researchers have compared trait accuracy at different levels of relationship. For example, cross-sectional studies have almost invariably shown that friends who have known the target for a considerable length of time agree with each other (consensus) and with the target's self-judgements (accuracy) more than do acquaintances (Kenny & Kashy, 1994) and relative strangers who have only briefly observed or interacted with the target (e.g.,

Cloyd, 1977; Funder & Colvin, 1988; Paunonen, 1989; Taft, 1955; Watson & Clark, 1991). Moreover, this acquaintanceship effect has remained robust when a number of obvious confounds (such as actual similarity) have either been experimentally or statistically controlled for. Hence, these results suggest that the beneficial effects of acquaintanceship on accuracy and consensus emerge more clearly in reasonably close and longstanding relationships, situations in which friends have had abundant opportunities to observe and interact with each other in different situations that are likely to be diagnostic of underlying personality traits.

Scant attention, however, has been directed at the highly intimate end of the relationship continuum, such as close romantic relationships. Thus, it is not clear whether trait accuracy continues to improve in the more intimate context of romantic relationships or whether it plateaus at the level of platonic friendship. To my knowledge, the only study that has explicitly compared the accuracy of judges in different close relationships found no difference between the level of average self-other agreement achieved by parents and college friends when judging the same target's personality traits (Funder, Kolar, & Blackman, 1995). A study by Macrae (1982) has sometimes been cited as evidence that spouses are better judges of each other's personality than people who are in different types of judge-target relationships. In this research, married couples attained an average self-spouse agreement correlation of about .59, considerably higher than the level of self-peer agreement typically achieved by friends in previous studies (about .30 to .40). However, Macrae (1982) used a simple dyadic design where each spouse only judged a single target (his or her own partner). As mentioned previously, this design is unable to cleanly separate the effect of the judge-target relationship from other characteristics of judges and targets. Moreover, Macrae (1982) also failed to control for the possible influence of assumed similarity on spouses' trait accuracy.

The Present Study

The current cross-sectional study compared the trait accuracy and consensus achieved by perceivers when judging targets at three different levels of acquaintanceship: dating partners, friends, and strangers. As previously noted, the data was gathered in the context of the first study which investigated empathic accuracy, but generated prior to the problem-solving discussions. Moreover, I sought to deal with a number of possible

confounds that could plausibly account for an acquaintanceship effect. As mentioned previously, one problem with many cross-sectional studies (including the current study) is that the use of nested designs, in which each perceiver judges a single target, fails to cleanly separate the effect of the relationship from the characteristics of the judge and target (Funder & Colvin, 1988; Kenny, 1994). I dealt with this problem in two ways. First, at each level of judge-target relationship, perceivers judged the same target. Hence, individual differences in target judgeability could not account for the presence of an acquaintanceship effect. Second, I measured and subsequently controlled for a number of characteristics of the judge that could plausibly covary both with levels of acquaintanceship and accuracy (e.g., relationship status, attributional complexity, various personality traits, etc).

In addition, due care was taken to attend to certain measurement artifacts identified by Cronbach (as outlined previously). The two elevation components, which refer to the possible influence of response styles shared by the judge and target, were dealt with by the use of correlational measures (Bernieri, Zuckerman, Koestner, & Rosenthal, 1994; Paunonen, 1989). Another possible artifact, assumed similarity, refers to judges projecting their own personality traits onto the target, and fortuitously, the judge and target happen to share similar traits. A number of researchers have argued that well-acquainted judges both tend to assume more similarity and to actually be more similar to the target in terms on a wide variety of characteristics than judges who are relative strangers (e.g., Kenny, 1994; but c/f. Funder et al., 1995). Hence, it is important to rule out the possible influence of actual (and assumed) similarity when assessing the link between acquaintanceship and trait accuracy. Stereotype accuracy, the final confound mentioned by Cronbach, can also artifactually inflate agreement correlations if the judge relies on his or her stereotype of the average person when judging the target's personality, and coincidentally, the target resembles the prototypical person. In the current study, the use by judges of both assumed similarity and stereotype endorsement were measured and statistically controlled for.

Note, however, that the way stereotype accuracy is measured in this study is distinct from the method typically used in previous research. When dealing with profile agreement scores, accuracy researchers have generally calculated a partial correlation between each set of acquaintance's and self-judgements, across traits, correcting for both

the average self-description and the average acquaintance's description (see Funder & Colvin, 1997, for more details). One problem with this procedure is that it does not necessarily measure judge-specific stereotypical beliefs. For example, a judge's subjective perception of the average male's or female's personality profile may be different from the objective or average stereotype measured across all perceivers and targets in the sample. Hence, the corrected accuracy scores may not control for each judge's particular response set or personally held stereotype. An alternative method employed in the current study, was to directly measure judge-specific stereotypes by simply asking each judge to provide ratings of the average person's personality, of a similar age and gender to the target (e.g., Gage, 1952; Sillars et al., 1984). These stereotype ratings were then correlated with the judge's ratings of the target's personality to form an index of stereotype accuracy. The effect of stereotype accuracy was subsequently controlled for when measuring the relationship between acquaintanceship and accuracy.

Predictions

Although existing research provides little guidance with respect to the likely performance of dating partners relative to friends, a prediction can be advanced based on the nature of these different types of relationships. Both dating partners and friends are likely to possess rich and complex knowledge structures concerning the personalities of their partners and friends, based on extensive observation and interaction with the target across a variety of settings. However, those in romantic relationships should have the edge over those in platonic relationships in terms of the acuity and diagnosticity of the underlying knowledge base. Compared to platonic relationships, romantic relationships typically involve: a) more time spent with the other person (Berscheid, Snyder, & Omoto, 1989), b) more open and intimate communication about private hopes, fears, wishes, and so forth (Levinger, 1980; Reis & Shaver, 1988), and c) more elaborate and complex shared knowledge structures or theories concerning each other's personality (Kelley et al., 1983). Thus, it was expected that dating partners should forge more accurate and detailed target-specific theories than friends. In addition, in line with previous cross-sectional research, it was expected that both dating partners and friends would generate

more accurate judgements of the target's personality than strangers who have merely observed the target's behaviour during a 5-min videotaped dyadic interaction.

Admittedly, strangers' judgements based on thin slices of behaviour typically yield beyond chance levels of accuracy and consensus (e.g., Borkenau & Liebler, 1992, 1993a, 1993b). However, such data-driven judgements by strangers are likely to generate less accurate personality judgements than the theory-driven inferences of well-acquainted judges acquired over a long history of interaction.

Based on the theories of RAM and WAM, and the results of previous cross-sectional research, I expected that consensus would improve as a function of increased acquaintanceship under conditions of less than perfect overlap. In particular, it was predicted that consensus between two well-acquainted judges would be significantly higher than consensus between a well-acquainted judge and a stranger.

In sum, it was predicted that increased acquaintanceship would be associated with more accurate trait judgements. Specifically, it was hypothesised that dating partners would be more accurate than friends, and both types of well-acquainted judges would be more accurate than strangers. In terms of consensus it was expected that judges who were mutually well-acquainted with the target (dating partners and friends) would generate a higher level of consensus when judging the targets' personality than that obtained by a well acquainted judge (dating partner or friend) and a stranger.

The Good Trait: The Influence of Trait Visibility

Some traits have certain properties that make them easier to judge than others. According to RAM, such judgeability is a function of differential relevance and availability of behavioural cues (Funder, 1995). This theoretical prediction is bolstered by a number of studies showing that good traits are more visible or observable. For example, researchers have examined trait judgeability and found that some traits (e.g., talkative) are generally judged with more accuracy and consensus than other traits (e.g., self-insight), primarily because they are revealed more directly and frequently in overt social behaviour (e.g., Biernieri et al., 1994; Funder & Dobroth, 1987; Kenrick & Stringfield, 1980; John & Robins, 1993). Hence, I predicted that accuracy and consensus would increase as a function of increasing levels of trait observability.

However, it is important to clarify what this prediction actually entails. Previous research suggests that the nature of trait visibility can vary according to the level of acquaintanceship. On the one hand, John and Robins (1993) found that in terms of the Big-5 dimensions of personality, reasonably well acquainted judges rated traits in the extraversion domain (e.g., sociable and enthusiastic) as the most observable, and traits related to intellect (e.g., reflective and imaginative) as the least observable. As acquaintanceship develops it is not surprising that traits like sociable, which are more directly and frequently revealed during positive social interaction are considered to be more observable than traits like reflective which must be inferred on the basis of ambiguous and less overt cues such as contemplative facial expressions or distracted responses (Funder, 1995). On the other hand, trait visibility at very low levels of acquaintanceship has been shown to be highly dependent on the nature of the thin slice of behaviour observed by the judge. For example, getting acquainted interactions between strangers (typically used by accuracy researchers) yields behaviour that is directly relevant to traits defined by extraversion, whereas intellectual tasks generate cues that are particularly diagnostic of traits relating to intellect (see Kenny et al., 1994). In the present study, strangers observed two dating partners (targets) engage in a 5-min. debate on capital punishment, before making judgements of the targets' personality. Given that the interactants were highly acquainted and were involved in a structured intellectual task, it was expected that strangers would find traits relating to intellect more visible, and thus easier to judge, than the other trait domains of the Big-5 (including extraversion). For the well acquainted judges in this study, it was expected that they would more accurately and consensually judge the most observable traits (those relating to extraversion) than the least observable traits (those relating to intellect).

Trait \times Information: The Interaction between Observability and Acquaintanceship

According to RAM, particular traits can only be inferred on the basis of certain kinds of information. This specific interaction between trait judgeability and information quality is referred to in Funder's model (1995) as *diagnosticity*. This prediction is supported by some evidence that the influence of trait visibility on trait accuracy is not

invariant. For example, Paunonen (1989) found that trait observability interacted with acquaintanceship in its effect on accuracy (but c/f. Funder & Colvin, 1988). At low levels of acquaintanceship, trait visibility was a strong determinant of accuracy, whereas highly acquainted individuals judged each other accurately, irrespective of the observability of the trait domain being judged.

In line with previous research and theory, I predicted that the influence of trait observability on accuracy and consensus would decrease as acquaintanceship increased. In particular, strangers who had viewed the targets only by means of a brief videotaped debate were expected to accurately and consensually judge observable traits relating to intellect. In contrast, it was expected that the highly acquainted judges in this study, by virtue of their extensive and detailed knowledge of even the more subtle and private facets of the target's personality, would generate accurate and consensual judgements of all of the Big-5 trait domains, regardless of trait visibility.

To summarise, it was hypothesised that dating partners and friends would accurately and consensually judge both observable and unobservable traits, whereas strangers would only accurately and consensually judge observable traits, such as those relating to intellect.

Overview of Study 2

In this study, I tested for the moderating effect of acquaintanceship, trait visibility, and the interaction between acquaintanceship and trait observability, on trait accuracy and consensus. One hundred dating partners (50 couples) served as both judges and targets, whereas 50 nominated friends of the dating couples and 50 strangers acted solely as judges. Dating partners provided self-descriptions, target-descriptions (their partner), and stereotypical-descriptions (the prototypical person of the same sex and age-range as the target person) on a version of the Big-5 personality scale. Similarly, friends provided self, target (their male and female friend), and stereotypical personality descriptions. The targets (dating couples) were then videotaped having a 5-min. dyadic debate on the use of capital punishment. In addition, strangers provided self and stereotypical personality descriptions, and after observing a videotaped debate by two targets on capital punishment, generated descriptions of the male and female target's personality. These

data were subsequently analysed to yield profile and trait correlations measuring accuracy (self-other agreement), consensus, assumed similarity, and stereotype accuracy.

To summarise the predictions, I hypothesised (a) that accuracy and consensus would increase with increased levels of acquaintanceship, (b) that accuracy and consensus would increase with increasing levels of trait visibility, and (c) that acquaintanceship and trait observability would interact in their effects on accuracy and consensus, such that the influence of trait observability on accuracy and consensus would decrease as acquaintanceship increased.

Results

Before describing the results, it is important to note the way interjudge agreement was calculated. Previous research has typically used correlation coefficients computed either across targets (commonly referred to as *trait correlations* or nomothetic analysis) or within targets or target-judge pairs (commonly referred to as *profile correlations* or idiographic analysis). Due to the nature of the questions being assessed in the present study, I have used both types of correlation to measure interjudge agreement. In the first two parts of the results section, I use profile correlations to present descriptive results and to test for acquaintanceship and sex differences in interjudge agreement. In the final two part of the results, I use trait correlations to test the influence of trait observability on interjudge agreement and to examine whether trait observability and acquaintanceship interact in their effects on trait accuracy and consensus. In all analyses, the target is the dating partner.

Descriptive Analyses

Interjudge agreement was examined in two ways. The first was to assess self-other agreement: the extent to which the target's and another person's view of him or herself are in agreement. This was done by computing profile correlations between the 20-item Big-5 personality ratings for three pairs of variables: the target and his or her dating partner, the target and his or her friend, and the target and the stranger. The resulting three types of correlation coefficients were subsequently treated as self-other

agreement scores. This procedure was repeated separately for each target (i.e. the male and female dating partners), resulting in six correlational indices.

The second way I examined interjudge agreement was to measure consensus: the degree to which judgements by different observers of the same target are in agreement. Consensus was derived in a manner similar to self-other agreement. In this case, however, profile correlations were computed between the 20-item Big-5 personality profiles for the same target provided by the following pairs of judges: the dating partner and the friend, the dating partner and the stranger, and the friend and the stranger. Once again, the resulting correlation coefficient was treated as a consensus score, and this procedure was repeated for each target.¹¹

The means and standard deviations for the different kinds of self-other agreement and consensus, derived from profile correlations, are presented in Table 13.¹² As predicted, even the agreement and consensus mean profile scores that include stranger's judgements of the target's personality are significant at above chance levels (i.e. at the .05 level, one tailed). Note that the various indices of interjudge agreement exhibit good variance.

¹¹ There was no correlation between male and female target scores on the various indices of interjudge agreement (using profile scores). Hence, all subsequent analyses are conducted with the male and female targets combined ($N = 100$)

¹² All within-target correlations were transformed via Fisher r to z formula before being analysed further. However when reporting the results, all mean scores were transformed back from Fisher's z , for ease of readability.

Table 13
Descriptive Statistics for All Interjudge Agreement Indices, Derived From Profile Correlations

Interjudge Agreement Indices	<i>M</i>	<i>SD</i>
Self-other agreement		
Self-partner	.56*	.28
Self-friend	.48*	.29
Self-stranger	.20*	.29
Consensus		
Partner-friend	.50*	.29
Partner-stranger	.18*	.31
Friend-stranger	.19*	.30

Note. $N = 100$. All correlations presented are converted back from fisher's z to improve readability.

* $p < .05$, one tailed.

Self-Other Agreement: Acquaintanceship and Sex Differences

To test whether self-other agreement increased with higher levels of acquaintanceship, and in particular, whether dating partners attained higher levels of agreement than friends, a 3 (stranger vs. friend vs. dating partner) \times 2 (female judge vs. male judge) between subjects analysis of variance was calculated with self-other agreement (profile) scores as the dependent variable.¹³ As predicted, the analysis revealed a significant main effect for acquaintanceship $F(2, 294) = 60.76, p < .001$. A Tukey post hoc test demonstrated that dating partners attained significantly higher levels of self-other agreement than both friends and strangers (see Table 13). The mean

¹³ A check on the independence of the various indices of self-other agreement revealed that some scores were correlated with each other, whereas other scores were unrelated. A similar pattern of dependence and independence was found within the consensus scores. Given this mixed pattern, a between subject ANOVA, rather than a within subject ANOVA, was used – a conservative analysis when dealing with within groups.

In addition, sex of the target could not be included as an independent variable in this analysis because for dating partners it was confounded with sex of the perceiver. However, a 2 (stranger vs. friend) \times 2 (female target vs. male target) between subjects analysis of variance was calculated to determine whether self-other agreement differed as a function of the sex of the target. The analysis revealed a significant sex of target effect $F(1, 196) = 19.17, p < .001$, with perceivers attaining higher levels of self-other agreement with female targets ($M = .45$), than with male targets ($M = .28$). In other words, women's personality traits seem to be more judgeable than men's. As expected, a main effect for acquaintanceship was also found, with greater self-other agreement produced by friends ($M = .48$) than by strangers ($M = .20$). There was no hint of an interaction effect ($F < 1$)

difference between friends and strangers was also significant. In contrast, there was no hint of a main effect for sex of judge $F(1, 294) < 1, n.s.$, or an interaction between acquaintanceship and sex of the judge, $F(2, 294) < 1, n.s.$

Alternative Explanations for the Acquaintanceship Effect

Before reaching the conclusion that this effect is the result of the differing levels of acquaintanceship between the judge and the target, a number of alternative explanations need to be considered. For all three levels of acquaintanceship, perceiver's judged the same target and therefore individual differences in target judgeability cannot account for the effect. However, there are certain characteristics of the judge, as well as a number of other variables, that could plausibly covary both with levels of acquaintanceship and self-other agreement, and hence could explain the acquaintanceship effect. These are dispositional characteristics, relationship status, assumed similarity, and shared stereotypes. I will consider each of these in turn.

Dispositional characteristics. Table 14 presents a profile of each class of judge in terms of their dispositional characteristics. As can be seen, self-reported mean levels of extraversion, conscientiousness, intellect, and agreeableness varied significantly across the different levels of acquaintanceship. To test whether such differences accounted for the acquaintanceship effect, I repeated the previous ANOVA, but covaried out the extraversion, conscientiousness, intellect, and agreeableness scores for each class of perceiver. The critical acquaintanceship effect (and the corresponding pattern of mean differences) remained reliable $F(2, 290) = 54.88, p < .001$. Hence, personality traits do not explain differing levels of accuracy at each level of acquaintanceship.

Table 14
Means of Dispositional Variables at Each Level of Acquaintanceship

Dispositional variables	<i>M</i>			<i>F</i>
	Dating partners	Friends	Strangers	
Attributional complexity	5.32	5.42	5.36	<1
Verbal intelligence	20.52	20.66	20.32	<1
Big-5				
Extraversion	6.47	6.09	6.13	2.57
Agreeableness	7.18	7.33	6.74	11.82
Conscientiousness	6.23	5.86	5.44	9.73
Emotional stability	5.23	4.97	5.04	1.68
Intellect	7.05	6.81	6.31	29.61

Note. All *F*s in boldface are significant at $p < .05$, two tailed.

Relationship status. Another artifactual explanation for the acquaintanceship effect is in terms of relationship status, which could also covary both with the level of acquaintanceship and self-other agreement. Indeed, not surprisingly, there was a much greater proportion of people who were currently in a romantic relationship in the dating sample (100%) than in the friend (46%) and stranger (66%) samples. It is conceivable that relationship participants, compared with those who are not in relationships, may inherently possess better social acuity or judgement ability. To test for the plausibility of such a claim, I carried out a 2 (stranger vs. friend) \times 2 (relationship non-participant vs. relationship participant) between subjects ANOVA with self-other agreement profile scores as the dependent variable.¹⁴ In accordance with the previous analysis, an acquaintanceship effect was found $F(1, 196) = 63.41, p < .001$. There was also no evidence of a main effect for relationship status $F(1, 196) < 1, n.s.$, or an interaction effect $F(1, 196) < 1, n.s.$ This result indicates that differences in perceiver's relationship status is an unlikely explanation for the acquaintanceship effect.

¹⁴ On this occasion an ANCOVA could not be used because all dating partners were relationship participants.

Assumed similarity. A third possibility is that people choose to become romantically involved with, and to a lesser degree befriend, others who are similar to themselves and, in turn, base their judgements of those others on their beliefs about their own personalities. Therefore, it is plausible that romantic partners outperform friends, and friends do better than strangers, only by virtue of self-selection and assumed similarity. To investigate this possibility, assumed similarity scores were derived by computing profile correlations for each judge (i.e. the dating partners, friends, and strangers), between their self and target ratings on the 20-item Big-5 personality inventory. The resulting correlation coefficients, representing the amount each judge assumed that the target's personality was similar to his or her own personality, was treated as an index of assumed similarity for each judge. This procedure was repeated separately for both targets (i.e. the male and female dating partners), resulting in six assumed similarity indices.

Next, the original ANOVA was recalculated, but with assumed similarity scores for each judge-target pair entered as a covariate. Consistent with the previous analysis, dating partners ($M = .51$) maintained greater levels of self-other agreement than friends ($M = .44$), whereas friends retained higher levels of self-other agreement than strangers ($M = .23$), $F(2, 293) = 51.85, p < .001$. This is a very similar pattern to the results previously obtained (see Table 12). This result implies that assumed similarity (along with self selection) cannot account for the acquaintanceship effect.

Stereotype accuracy. A fourth possible artifact is that self-other agreement could be spuriously inflated by a stereotype component pervading the judgements. To illustrate, a judge's other rating (of a female target's personality) could be correlated substantially with the target's self rating if (a) the judge guessed by using his or her stereotype response for each item (i.e. the average female's response), and (b) the target's self ratings actually resembled the judge's ratings of the prototypical female's personality (Paunonen, 1989). This process, often referred to as stereotype accuracy, can render profile agreement scores essentially uninterpretable (Cronbach, 1955; Kenny, 1993; but see Jackson, 1982). To control for the influence of judges' stereotypes on the acquaintanceship effect stereotype accuracy scores were generated in the same fashion as for assumed similarity. In this case, however, profile correlations were computed for each judge (i.e. the dating partners, friends, strangers) between their gender-specific

stereotype ratings and their ratings of the target on the 20 item Big-5 personality inventory. The resultant correlation coefficients were treated as an index of stereotype accuracy for each judge, and this procedure was subsequently repeated for both targets (the male and female dating partners). These scores were then entered as a covariate in the original ANOVA. The critical acquaintanceship effect remained robust $F(2, 293) = 56.09, p < .001$. Moreover, the means obtained were very similar to the previous analysis: dating partners ($M = .55$), friends ($M = .49$), and strangers ($M = .21$). Hence, stereotype accuracy cannot explain the acquaintanceship effect.

To summarise, as predicted self-other agreement increased with greater levels of acquaintanceship. In particular, dating partners attained significantly higher levels of self-other agreement than friends, whereas friends generated reliably higher levels of self-other agreement than strangers. Furthermore, the acquaintanceship effect remained robust when plausible alternative explanations such as the influence of judges' dispositional characteristics, relationship status, assumed similarity, and stereotype accuracy were controlled for.

Consensus and Acquaintanceship

To test whether consensus varied as a function of the level of acquaintanceship, a one-way (friend-stranger \times partner-stranger \times partner-friend) ANOVA was performed with profile consensus scores as the dependent measure. As expected, a significant acquaintanceship effect was attained $F(2, 297) = 48.41, p < .001$. Post-hoc tests showed that consensus between judges who were both well acquainted with the target (i.e. dating partners and friends, $M = .50$) was significantly higher than consensus attained by dating partners and strangers ($M = .18$), and friends and strangers ($M = .19$).

Differences Between Traits: The Influence of Trait Observability

Recall that observable traits were expected to be more easily judged than unobservable traits. This prediction was examined in two ways. The first was to assess self-other agreement by calculating *trait* correlations between the Big-5 factor scores (i.e. extraversion, agreeableness, conscientiousness, emotional stability, and intellect) for

three pairs of variables: the target and his or her dating partner, the target and his or her friend, and the target and the stranger. The resultant correlation coefficients were treated as an index of self-other agreement for each type of judge on the five personality factors. This procedure was repeated for both targets (the male and female dating partners). The second way was to compute consensus scores by calculating trait correlations on the Big-5 factor scores for the same target provided by the following three pairs of judges: the dating partner and the friend, the dating partner and the stranger, and the friends and the stranger. Once again, the resulting correlation coefficient was treated as an index of consensus for each pair of judges on the Big-5 factor scores, and this procedure was repeated for both targets (the male and female dating partners).¹⁵

The correlational analyses reveal that accuracy and consensus improved as a function of increasing levels of trait visibility, regardless of the level of acquaintanceship between the judge and the target. Specifically, dating partners and friends attained greater self-other agreement and consensus for the more observable traits in the extraversion domain than the less observable traits in the intellect domain (see Table 15). In general, strangers also achieved significant, albeit modest levels of self-other agreement and consensus for more visible traits related to intellect (due to the intellectual nature of the 5-min. interaction they observed), but not for the less visible traits related to other dimensions of the Big-5. Only 2 of 15 agreement and consensus correlations involving strangers' judgements of each target's personality deviated from this pattern. That is, consensus between friends and strangers was obtained for emotional stability rather than for intellect.

Note that the pattern of mean self-other agreement and consensus *trait* correlations for each class of judge (see the bottom row of Table 15), is very similar to the previously described acquaintanceship effect derived from *profile* correlations shown in Table 13. In both cases, self-partner agreement was higher than self-friend agreement, whereas self-friend agreement was greater than self-stranger agreement. Similarly, consensus between partners and friends was higher than consensus between partners and strangers and consensus between friends and strangers for both the trait and profile

¹⁵ One advantage to using trait correlations is that they are completely immune to the artifactual influence of assumed similarity and stereotype accuracy (see Funder & Colvin, 1997).

correlations. Hence, regardless of how self-other agreement and consensus was calculated the acquaintanceship effect is robust.

Table 15
Correlations Between Big-Five Factors and Interjudge Agreement Indices, Calculated Across Participants

Big-5 Factors	Self-other agreement			Consensus		
	Self- partner	Self- friend.	Self- stranger	Partner- friend	Partner- stranger	Friend- stranger
Extraversion	.48*	.45*	.11	.39*	.04	.00
Agreeableness	.39*	.20*	.06	.31*	.04	.10
Conscien- tiousness	.44*	.36*	.00	.34*	.12	.11
Emotional stability	.35*	.43*	.13	.42*	.00	.21*
Intellect	.29*	.20*	.20*	.20*	.23*	.04
Average <i>r</i>	.41*	.34*	.10	.34*	.09	.09

Note. *N* = 100. All correlations presented are converted back from fisher's *z* to improve readability.

* *p* < .05. two tailed.

Interjudge Agreement and the Interaction of Trait Visibility and Acquaintanceship

Table 15 reveals a marked difference in the level of self-other agreement and consensus attained for the various personality factors by judges who were well acquainted with the target (i.e., dating partners and friends), compared with those who were strangers. Partners and friends demonstrated significant agreement and consensus on all five personality factors, regardless of trait observability. In contrast, strangers attained significant self-stranger agreement and partner-stranger consensus only for observable traits relating to intellect. Hence, observability and acquaintanceship interacted in their effects on interjudge agreement, such that the influence of trait

observability on accuracy and consensus decreased as the level of relationship between the judge and the target increased.

Discussion

Summary of Findings

I proposed three hypotheses in this study, based on the derivations of RAM (and relevant prior research). All three predictions were confirmed by the results of this study. First, both trait accuracy and consensus improved as a function of increased acquaintanceship. Specifically, dating partners were more accurate than were friends, and both kinds of well-acquainted judges were more accurate than were strangers. Moreover, the acquaintanceship effect remained robust when several alternative explanations were eliminated or controlled for.

Second, accuracy and consensus improved as a function of increased levels of trait behaviour observability. That is, dating partners and friends found it easier to judge traits in the extraversion domain than the intellect domain. Similarly, strangers attained a modest, yet significant level of accuracy and consensus on the more visible traits relating to intellect (I argue because of the intellectual nature of the brief interaction they observed), but not on the less visible traits relating to the other dimensions of the Big-5.

Finally, acquaintanceship and observability interacted in their effects on interjudge agreement, such that the influence of trait visibility on trait accuracy and consensus decreased as acquaintanceship increased. Dating partners and friends accurately and consensually judged both observable and unobservable traits (all of the Big-5 trait domains), whereas strangers accurately and consensually judged only observable traits relating to intellect.

Acquaintanceship, Trait Accuracy, and Consensus

The basic finding that the more information the better judges performed corresponds with the dictates of both conventional wisdom and a good deal of prior research and theorising. A plethora of cross-sectional studies have documented that well-acquainted judges attain more accurate and consensual personality judgements than strangers who have only briefly observed or interacted with the target (e.g., Funder &

Colvin, 1988; Paunonen, 1989; Taft, 1955; Watson & Clark, 1991). From the perspective of RAM (Funder, 1995), as acquaintanceship increases judges gain access to greater quality and quantity of information about the target ("good information") which results in more accurate and consensual judgements of the targets' personality.

Clearly, my results contradict the general argument that accuracy plateaus at very low levels of acquaintanceship (e.g., Ambady & Rosenthal, 1992; McArthur & Baron, 1983, Wilson & Schooler, 1991). However, it was not the case that strangers' personality judgements were completely inaccurate. In accordance with previous research (e.g., Borkenau & Liebler, 1992, 1993a) strangers achieved a reliable, albeit modest, degree of accuracy and consensus (based on profile analyses). Moreover, this effect remained robust when strangers' use of assumed similarity and stereotype endorsement was statistically controlled for. That is, it appears that strangers in this study based their personality judgements on the behavioural cues evinced by the target during the interaction rather than on their pre-existing response-set biases.

A striking feature of the results was that dating partners outperformed friends. Although both kinds of well-acquainted judges demonstrated considerable understanding of the target's personality, it is likely that dating partners possessed a more detailed and valid knowledge base and personality theory than friends because: (a) they have more opportunities to observe and interact with the target over a wide range of situations, and (b) they engage in more open and direct communication that reveals subtle and private facets of the target's personality. Indeed, disclosure concerning targets' inner thoughts and feelings has been shown to be particularly diagnostic of underlying personality traits (Anderson, 1984). The only other study to compare judges' performance across different close relationships, found that parents were no more accurate than college friends when judging the target's personality characteristics (Funder et al., 1995). Thus, the present findings represent the first demonstration (that I am aware of) that trait accuracy does not plateau at the level of platonic relationships, but continues to improve in the more intimate context of romantic relationships.

Alternative explanations. It is important to note that clear evidence for the acquaintanceship effect was found regardless of the method used to measure trait accuracy and consensus (i.e. using either within-participant profile scores or across-participant trait scores). Moreover, my analyses and research design ruled out several

plausible competing explanations for the link between acquaintanceship and trait accuracy. Judges who were better acquainted with the target did not generate more accurate personality judgements because they (a) perceived targets who were easier to judge, (b) possessed certain personality traits that enhanced their judgement accuracy, (c) were currently involved in romantic relationships, (d) assumed higher levels of similarity, or (e) made greater use of stereotypes.

Of particular interest is the finding that judges' use of assumed similarity did not account for the acquaintanceship effect. Several researchers have speculated that the acquaintanceship effect may be artifactually inflated by well-acquainted judges both tending to assume more similarity and to actually *be* more similar to the target on a wide range of psychological attributes, compared to judges who are relative strangers (e.g., Kenny et al., 1994). This study found no evidence for this claim. Well-acquainted judges did not project their own personality traits onto the target to a greater extent than did strangers. In fact, dating partners and friends were actually no more similar to the target, in terms of their personality traits, than were strangers (see Funder et al., 1995).

The results relating to interjudge consensus were also generally consistent with prior theory. As predicted by both RAM (Funder, 1995) and WAM (Kenny, 1991, 1994), consensus improved as a function of increased acquaintance. Recall that WAM identifies three possible explanations for the acquaintanceship effect (in terms of consensus): overlap, communication, and shared meaning systems. Although it is not possible to directly test these explanations in the present study, the shared meaning system explanation is undermined by the fact that acquaintanceship was unrelated to the amount of actual similarity between judges' personality traits. Moreover, the few studies that have tested the overlap and communication parameters of Kenny's model (under conditions of less than perfect overlap) have not found them to be necessary conditions for the acquaintanceship effect (e.g., Funder et al., 1995).

Probably the most parsimonious explanation for the positive link between acquaintance and consensus found in this study is the informational hypothesis posited by RAM. That is, relationship targets make available more relevant information of their own personality to their partner or friend, and thus the chances of the judge detecting and relying on such information improves. Although, WAM's emphasis on the amount of information explicitly shared between judges may sometimes be important, particularly

under conditions of perfect information overlap (Blackman & Funder, 1998), in my view the primary determinant of the acquaintanceship effect is likely to be “good information”.

Trait Observability, Trait Accuracy, and Consensus

This study did not directly and independently assess trait observability. Nevertheless, it is reasonable to expect that the traits I measured possessed the same variability in observability as reliably shown in prior research (e.g., John & Robins, 1993; Kenny et al., 1994). Thus, I used this prior evidence to make *a priori* classifications of which traits would be high or low in terms of observability. As expected, this study found evidence that trait observability uniquely moderated trait accuracy and consensus. Judges inferred observable traits more easily than unobservable traits, regardless of their level of acquaintanceship with the target.

From RAM’s vantage point (Funder, 1995), the positive link between trait visibility and accuracy must be explained in terms of differential relevance and availability of trait-related behaviours. In the present study strangers who observed couples in romantic relationships participate in a brief debate accurately and consensually judged only observable traits- namely those related to intellect. Applying the logic of Funder’s model, it is not surprising that in a highly structured task, in which two romantic partners are displaying their intellectual wares and mounting arguments and counter-arguments, that traits related to intellect are relatively visible, and thus easier to judge, than other trait domains (including extraversion). Indeed, previous research suggests that intellectual tasks generate behavioural cues that are especially relevant to traits defined by intellect, whereas the typical unstructured “getting-acquainted” interaction between strangers yield cues that are particularly diagnostic of traits relating to extraversion (see Kenny et al., 1994).

At first glance, the finding that well acquainted perceivers in this study found traits related to extraversion the easiest to judge and traits in the intellect domain the most difficult to judge (the opposite pattern of results to strangers) does not support the predicted main effect of trait observability. However, recall that I expected the nature of trait observability to vary as a function of acquaintanceship, because of its malleability at very low levels of acquaintanceship. Personality judgements based on thin slices of behaviour are exquisitely sensitive to the nature of the particular interaction. Depending

on the context, certain behavioural cues may be abnormally stifled while others may be uncharacteristically overt or even exaggerated, resulting in an idiosyncratic display of trait behaviours. In contrast, at higher levels of acquaintanceship in which judges have observed and interacted with the target across a much broader range of situations, perceivers are likely to form more reliable and enduring impressions of the most salient features of the targets' personality.

Acquaintanceship \times Observability

The final aim of the present study was to determine whether observability and acquaintanceship interact in their effects on accuracy and consensus. As predicted, the influence of trait observability on accuracy and consensus decreased as the level of acquaintanceship between the judge and the target increased. In particular, as discussed previously, it is hardly surprising that unobservable traits are not accurately rated by strangers, because of the lack of available and relevant behavioural cues (Funder, 1995). Remarkably, however, all of the Big-5 domains of personality were judged quite accurately by the well-acquainted partners and friends, regardless of trait observability.

Previous research has typically shown that while increased acquaintanceship leads to more accurate judgements of observable traits, unobservable trait domains continue to be poorly evaluated (e.g., with interjudge agreement of .10 to .20 being common). Why were such comparatively high accurate judgements attained in the current study? The most plausible explanation is that the present study used romantic partners and friends who were considerably more intimately acquainted with the target (mean relationship length close to 7 years) than the roommates and acquaintances predominantly used in prior research. At such deep levels of acquaintanceship, it seems possible for judges to attain quite respectable levels of accuracy and consensus for even the more latent and covert personality traits.

STUDY 3

THE LINK BETWEEN EMPATHIC AND TRAIT ACCURACY: THE QUEST FOR CONSISTENCY IN JUDGEMENT ACCURACY ACROSS DOMAINS

Throughout the history of accuracy research, the existence of the good judge has remained intuitively appealing, yet frustratingly elusive. However, a number of recent studies have yielded fruitful results. Reliable individual differences in judge's empathic ability have been demonstrated both across time (Thomas & Fletcher, 1997) and across different relationships or interactions (Study 1 of this thesis). Moreover, Vogt and Colvin (1998) revealed cross-target consistency by judges when describing the personality attributes of strangers. Hence, some evidence has accumulated for the existence of individual differences in judge's ability *within* particular judgement domains. Nevertheless, it could be argued that a further test of the good judge is the generality of performance *across* judgement domains. For example, is the good judge of personality also a good judge of people's cognitive and affective states? Although, no prior research has addressed this important question, there are a number of general arguments both for and against the concept of cross-domain consistency in judgement accuracy.

On the one hand, the primary characteristic of social intelligence, as typically defined, is the general ability of perceivers to judge the internal psychological states and trait of others (see Sternberg & Smith, 1985). In fact, Ickes (1993) has claimed that the accurate judgement of states may well be a necessary, if not sufficient, condition for the accurate judgement of traits. As Ickes states, "how can I know that you are consistently morose unless I know that you are in a bad mood today, just as you were yesterday?" (Ickes, 1993, p. 587). Indeed, there is some empirical evidence that indirectly supports this argument. Anderson (1984) found that listening to a target disclose his or her thoughts and feelings lead to more trait accuracy than when listening to the same person describe his or her hobbies and activities.

On the other hand, Kenny and Albright (1987) in their review of the accuracy literature concluded that the available evidence for generalised (or consistent) accuracy was weak. Moreover, the correlates of judgement ability have not formed a reliable and coherent pattern, varying both across relationships (Study 1 of this thesis) and across

judgement domains (Schneider, et al., 1979). These kinds of inconsistent results have lead some researchers to conceptualise social intelligence in terms of highly domain-specific abilities (e.g., Cantor & Kihlstrom, 1987).

In this final study, I examined judges' accuracy within targets but across judgement domains. Specifically, judges' degree of accuracy and consensus in describing the targets' personality profile were correlated with their accuracy in judging the same targets' thoughts and feelings.¹⁶ The question here is whether people who are good at making empathic judgements are also good at making personality judgements (of the same targets). These correlations were performed at three different levels of acquaintanceship – dating partners, friends, and strangers.

One way to conceptualise these analyses is in terms of an important, previously made distinction – data-driven versus theory-driven accuracy (Thomas & Fletcher, 1997). The results of this thesis and other research have demonstrated that strangers routinely rely on the behavioural information evinced by the target when generating empathic and trait judgements. Hence, if some targets provide data that is more diagnostic of *both* their underlying psychological states and traits than others, then consistent levels of data-driven accuracy should be attained.¹⁷ In other words, for consistent accuracy (across domains) to be achieved by strangers, individual differences in target's readability should also extend across judgement domains (the quintessential good target). In contrast, well-acquainted judges are likely to rely to an important degree on their theories when deriving their judgements (Thomas & Fletcher, 1997). Thus for consistent theory-driven accuracy to occur, for judges who are friends or romantic partners, then it follows that individuals should essentially be using the same theories to drive both their personality and empathic judgements. Those who have good general models or person-specific theories should do well across both domains, whereas those who have weak theories should do poorly. On the other hand, if judgements across the two domains (empathic

¹⁶ As mentioned previously, there are two ways of measuring accuracy and consensus of personality judgements – profile scores and trait/item scores. When examining individual differences, profile scores need to be used because they measure the accuracy and consensus attributable to a particular judge-target pair for the target's entire set of personality characteristics. Empathic accuracy is also invariably measured idiographically.

¹⁷ It is important to note that in this thesis strangers' personality and empathic judgements were generated in the context of different interactions. Personality judgements were made after observing the targets debating the use of capital punishment, whereas empathic judgements were generated from targets' problem-solving interactions.

and trait accuracy) call upon very different knowledge structures, then one might predict a weak or null relation in performance across the two domains.

Given the lack of previous research and theoretical consistency involved, no specific predictions were advanced concerning the association between judges' accuracy (and consensus) in describing the target's personality and their accuracy in judging that person's thoughts and feelings.

Results

The question addressed in this analysis was the extent to which perceiver's empathic accuracy covaried with their ability to judge the same target's personality. Pearson's correlations were computed for each type of perceiver between his or her empathic accuracy score and his or her self-other agreement and consensus profile scores, when judging the same target. The relevant data for the male and female dating partners is shown in Table 16. As can be seen, when judging their respective partner, none of the correlations between dating partner's empathic accuracy and the various kinds of self-partner agreement and consensus were significant. In fact, 3 of 6 correlations were negative.

Table 16
Zero-order Correlations of Male and Female Dating Partners' Empathic Accuracy Scores With Selected Interjudge Agreement Indices

Interjudge agreement index	Male partner EA	Female partner EA
Self-other agreement		
Self-male partner	-.08	---
Self-female partner	---	.16
Consensus		
Male partner-friend	-.09	---
Female partner-friend	---	.03
Male partner-stranger	.03	---
Female partner-stranger	---	-.14

Note. EA refers to empathic accuracy.

A similar pattern of results can be found in Table 17. All of the correlations between the level of self-other agreement and consensus attained by friends and strangers and their associated empathic accuracy scores for the same target were non-significant. Indeed, they were all close to zero, and 6 of the 12 correlations were negative. Overall, these results reveal no evidence for general individual differences in people’s ability to judge both a target’s personality (when measured by the Big-5 factor model) and transient on-line cognitions and emotions.

Table 17
Zero-order Correlations of Selected Interjudge Agreement Indices With Friends’ and Strangers’ Empathic Accuracy Scores for Male and Female Dating Partners

Interjudge agreement index	Friend EA for		Stranger EA for	
	Male Partner	Female Partner	Male Partner	Female Partner
Self-other agreement				
Male self-friend	-.18	---	---	---
Female self-friend	---	-.15	---	---
Male self-stranger	---	---	.07	---
Female self-stranger	---	---	---	.00
Consensus				
Male partner-friend	-.15	---	---	---
Female partner-friend	---	-.17	---	---
Male partner-stranger	---	---	-.10	---
Female partner-stranger	---	---	---	.16
(Male) friend-stranger	.03	---	-.18	---
(Female) friend-stranger	---	.07	---	.11

Note. EA refers to empathic accuracy.

Discussion

Consistency in Judgement Performance Across Trait and Empathic Accuracy Domains

Recall that this final study examined whether people who are good at making empathic judgements are also good at making personality judgements (of the same targets). This question was addressed within three levels of acquaintanceship (dating partners, friends, and strangers). The results revealed no association between judges' accuracy (and consensus) in describing the target's personality traits and their accuracy in judging that person's thoughts and feelings.

This set of findings can be interpreted in terms of data-driven versus theory-driven accuracy (Thomas & Fletcher, 1997). There is good evidence that strangers' judgements (both empathic and trait) are primarily data-driven, whereas dating partners and friends tend to rely on their theories in deriving their judgements. Hence, the finding that strangers did not attain consistent data-driven accuracy across judgement domains, suggests that targets were not consistent in the degree to which they evinced behavioural cues that were diagnostic of both their underlying personality traits and their cognitive and affective states. In short, there appears to be no evidence for the existence of the good target across the empathic and trait accuracy domains. In a similar vein, the failure of well-acquainted judges to generate consistent theory-driven performance across judgement domains suggests that both dating partners and friends are relying on different theories or knowledge structures when deriving their empathic and trait judgements.

This null result needs to be placed in the context of two prior studies which have found evidence of strong individual differences in ability operating within the empathic and personality domains (Marangoni et al., 1995; Vogt & Colvin, 1998). The finding that there are no significant relations across empathic and personality domains suggests that "social intelligence" may operate within distinct modules or domains to a much greater extent than is often supposed.

However, an important principle emanating from the attitude-behaviour consistency literature is that general attitudes predict general aggregates or patterns of behaviour, whereas specific attitudes predict specific kinds of behaviour (e.g., Aizen & Fishbein, 1977). Similar arguments have been raised by personality researchers

concerning cross-situational consistency in behaviour (e.g., Epstein & O'Brien, 1985; Kenrick & Funder, 1988) and behavioural prediction (Colvin & Funder, 1991). The point here is that the size of the relationship between two constructs (such as attitudes and behaviour) is dependent on the degree to which such constructs are similar in terms of their measurement specificity (sometimes referred to as bandwidth).

In the current study, the target's Big-5 personality attributes constitutes a general construct, the target's on-line thoughts and feelings experienced during a specific problem-solving interaction is a rather specific construct. Thus, we might expect to find low correlations between the accuracy obtained in the two domains by lay perceivers. If the bandwidth were adjusted to bring them more into line, we should find that related accuracy correlations increase in size. For example, accuracy in perceiving the target's on-line feelings or emotions (and not cognitions) may be reasonably positively related to accuracy in judging traits that have a strong affective or emotional tone (e.g., angry, moody, emotional). Alternatively, if empathic judgements were extended to more trait-like states (e.g., making judgements of an individual's general tendency across interactions to feel angry, defensive, think generous thoughts) the accuracy of such judgements may be closely related to Big-5 trait judgements.

Nevertheless, my finding of a null relation across the two judgement domains, does raise important questions about the concept and measurement of "social intelligence" viewed as a broad ability extending across every social judgement domain.

GENERAL DISCUSSION

In the general discussion I will first recapitulate the major results from each of the three studies of this thesis. The effect of individual differences in judgement ability on both empathic and trait accuracy will then be considered. Next, the central feature of this thesis will be reviewed – the major influence of the judge-target relationship on empathic and trait accuracy. Finally, I will raise several caveats concerning this research and discuss some general implications of these findings for the accuracy literature.

Summary of the Major Results

- (1) Greater acquaintanceship was associated with higher levels of both empathic accuracy and trait accuracy (and consensus).
- (2) Several relationship-level predictors of empathic accuracy were revealed with higher levels of relationship closeness, prior problem-specific disclosure, and shared cognitive focus being related to higher levels of empathic accuracy (although results varied according to the level of acquaintanceship).
- (3) No evidence for the good target of empathic judgements across three different levels of acquaintanceship was found.
- (4) Compelling evidence for individual differences in judges' empathic ability across different levels of acquaintanceship was demonstrated.
- (5) The gender of the judge was found to be an important predictor of generalised empathic ability (across different relationships or interactions), with females displaying greater judgement ability than males. However, other characteristics of the good judge, such as greater verbal intelligence and attributional complexity, were significantly moderated by the level of acquaintanceship between the judge and target.
- (6) Greater trait observability was related to higher levels of trait accuracy and consensus.
- (7) The influence of trait observability on trait accuracy and consensus decreased as the level of acquaintanceship between the judge and the target increased.

- (8) No evidence for individual differences in judgement ability across the trait and empathic accuracy domains was found.

At first glance, this summary may seem to reflect a potpourri of findings that vary both across and within judgement domains. However, RAM (Funder, 1995) provides a valuable theoretical framework for organising these results in both the empathic and trait accuracy domains. Recall, that Funder's model proposes four basic moderators that make accurate judgement more or less likely, two of which feature prominently in the results of this thesis – judge-based moderators (the good judge) and relationship-based moderators (the good relationship). In addition, Thomas and Fletcher's (1997) social cognitive approach can be used to explain the pivotal role played by the judge-target relationship in determining judgement accuracy. I next discuss these two moderators and what they tell us about judgement accuracy.

Individual Differences in Judgement Accuracy

The Case for the Good Judge

The concept of the good judge has intrigued psychologists since the inception of accuracy research. Nevertheless, weak results, combined with difficult methodological problems, have led most researchers to abandon the search entirely. However, the results of this study (and a few others) suggest that this reaction is premature. Recently, evidence for the good judge of underlying traits (Vogt & Colvin, 1998) and states (Marangoni et al., 1995) has been found when strangers judged targets. The current study provides an important extension of such research by demonstrating that individual differences in judges' empathic ability extend across different relationships or interactions. This novel finding reveals that the ability to make empathic judgements that are predominantly data-driven is closely allied to the ability to make such judgements when they are predominantly theory-driven. Taken together these results clearly support the existence of a strong "social intelligence" disposition, at least in the domain of empathic judgements.

Why does my research provide such compelling evidence for the good judge, when previous research has typically produced such a motley and inconsistent set of

findings? There appear to be at least three features that distinguish the studies that have found evidence for the good judge from those studies that have not. These features of successful studies are a) the use of multiple targets, b) the use of a challenging and extensive judgement task, and c) the use of a task that generates high levels of motivation. I discuss each feature in turn.

First, successful studies have used research designs where each perceiver judges multiple targets. This feature enables the effect of the judge to be disentangled from that of the target and the judge-target relationship.

Second, evidence for the good judge has been found in studies that have used a relatively difficult task to assess empathic accuracy, such as the self-revealing and affect-laden behaviour generated during therapist-client interactions (Marangoni et al., 1995) or problem-solving discussions between dating couples (the present study). These behaviourally rich and complex interactions provide judges with sufficient opportunity to showcase their empathic talents (for a related finding see Fletcher et al., 1992). In contrast, previous empathic accuracy research (mainly by Ickes and his colleagues) has most commonly used scripted and presumably often desultory getting-acquainted interactions between strangers who are being surreptitiously videotaped while ostensibly waiting for an experimental task to commence (see Ickes 1993 and Hancock & Ickes, 1996). Related empathic accuracy tasks are probably often either relatively easy or impossibly difficult. On the one hand, many of the empathic judgements are likely to be easy because strangers will be thinking and behaving in a relatively stereotypical fashion, or will be expressing their superficial thoughts and feelings in a transparent fashion. On the other hand, such unstructured interactions between two strangers are likely to be marked by frequent periods of silence, where no clues are evinced as to inner mental states. To reiterate a previous illustration, a very easy or an incredibly difficult exam is much less likely to discriminate between good and poor students than is a moderately difficult and challenging exam.

The final feature of studies that have found evidence for the existence of the good judge is that they have utilised tasks that concern intriguing topics - targets' personal and relationship problems- and are behaviourally rich. Thus, they are likely to be considerably more engaging and motivating for judges than are superficial and mundane interactions between strangers.

The Profile of the Good Judge

Identifying individual differences in judgement ability is one thing; determining what dispositions predict the good judge is quite another. This thesis has unearthed some of the first clues as to the psychological profile of the good judge of others' cognitive and affective states. Superior empathic ability was predicted by gender (women were better than men), higher verbal IQ, and higher attributional complexity. Interestingly, and lending further confidence to our findings, these three characteristics have also emerged as the most consistent (albeit modest) predictors of judgement ability in the trait accuracy literature (see Davis & Kraus, 1997; Vogt & Colvin, 1998).

However, the current research has provided perhaps the most rigorous test of the reliability of these individual difference markers to date, by testing whether they consistently predict judgement ability across markedly different levels of acquaintanceship. The results showed that both verbal intelligence and attributional complexity varied according to the nature of the judge-target relationship. At low levels of acquaintanceship, verbal intelligence was a good predictor of the good judge, whereas attributional complexity was a good predictor at high levels of acquaintanceship. Although this specific pattern of findings was unexpected, this pattern of findings is consistent with the overarching theme of this thesis that the sources of empathic accuracy vary as a function of the level of acquaintanceship. On the one hand, judges with higher levels of verbal intelligence are better suited at solving complex novel tasks such as detecting which behavioural cues are diagnostic of strangers' inner thoughts and feelings (i.e. data-driven accuracy). On the other hand, more attributionally complex judges appear to be more proficient in situations that require the integration of complex pre-existing theories with incoming behavioural data, such as making empathic judgements of well-acquainted targets during problem solving interactions.

Although the idea that the predictors of judgement ability may vary as a function of acquaintanceship is not new (Allport, 1937; Vernon, 1933), this research is the first to empirically confirm this notion. Furthermore, evidence for such a strong interaction provides one reason why previous attempts to identify the characteristics of the judge have generated such an inconclusive pattern of results.

Interestingly, the one variable that was consistently associated with judgement ability across all levels of acquaintanceship was the gender of the judge. Females were

consistently more accurate judges of others' thoughts and feelings than were males. It appears that there is more than just a kernel of truth to the conventional wisdom concerning women's superior social intuition and perceptiveness compared to that of men. Moreover, these results are consistent with the view that women are socialised from an early age to be more sensitive to, and more focused on relationships, and indeed other people in general, than are men (e.g., Acitelli & Young, 1996; Bakan, 1966; Cross & Madson, 1997). In short, women possess higher levels of interpersonal and relationship expertise than do men. This generalisation, however, begs the question as to the psychological mechanisms that underlie the link between gender differences in relationship-orientation and judgement ability. Applying the reasoning of Thomas & Fletcher (1997), there are a number of plausible sources of such gender differences in judgement accuracy.

First, it is probable that women are better (and perhaps more motivated) than men at detecting the complex behavioural data evinced by targets and integrating such cues in a relevant and accurate manner when making such data-driven judgements.

Second, women who are well acquainted with the target are likely to possess more elaborate and valid specific theories concerning the target/relationship than are men. Such detailed theories are likely to yield more accurate judgements because (a) they can be used to detect and utilise the subtle idiosyncratic cues evinced by the target that are diagnostic only to an insider, and (b) even in the absence of relevant behavioural cues, insiders can use their prior relevant knowledge and theories that are based on extensive prior observation and diagnostic discussion with the target.

The results of this thesis provide one important clue as to how women, as opposed to men, develop sophisticated and accurate theories of their own partner/relationship. The extent to which dating partners talked about the relationship problems with their close friend strongly predicted the empathic accuracy attained by female (but not male) dating partners. In a nutshell, the key difference between men and women seems to consist in the quality of men and women's problem-specific disclosure. Women seem to use close friends as a sounding board to vent their relatively raw and uncensored thoughts and feelings concerning the serious issues and problems in their romantic relationships. Such diagnostic disclosure is likely to lead to in-depth and frank discussions about the underlying causes of these problems and the respective feelings and intentions of both

parties, which in turn should broaden and hone women's theories and knowledge of their own relationship and partner. In contrast, men probably engage in fairly superficial and descriptive problem-related conversations with close friends that fail to generate the kind of feedback or insight that improves the validity of their pre-existing partner/relationship-specific theories. This gender difference in the quality of disclosure was also revealed by the influence of such discussions on the accuracy of the confidants' (or close friends') empathic judgements. Greater disclosure by female (but not male) dating partners led to superior empathic performance of the close friend when judging the dating couple.

Further evidence for the link between diagnostic disclosure by women and the effectiveness of their theory-driven judgements was shown by the evidence for the following mediational model: female (but not male) dating partners with higher levels of attributional complexity discussed more assiduously, and in greater depth, their relationship problems with their friends, which in turn produced greater levels of empathic accuracy when judging their dating partners. This general pattern of findings suggests that women utilise social support more effectively than men, by engaging in more diagnostic disclosure and in-depth discussion, to help supplement and refine their pre-existing working models of their romantic relationships, resulting in more accurate theory-guided judgements of their partner's thoughts and feelings.

Finally, women may possess more elaborate and intelligent general theories and stereotypes concerning close relationships (than do men), which in turn they use to more effectively guide their interpersonal judgements. Indeed, women have been found to use their stereotypes more effectively when making judgements of targets' personality traits than men (Vogt & Colvin, 1998). Note that it is a profound mistake to view this type of stereotype accuracy as merely a Cronbachian artifact, because women were highly selective rather than haphazard when using their stereotypes to guide their judgements (see Thomas & Fletcher, 1997; and Vogt & Colvin, 1998, for a more detailed discussion of the artifactual vs. substantive use of judges' general pre-existing knowledge structures).

The Breadth of Judgement Ability

Despite the fact that a similar pattern of gender differences were found in the empathic (this study) and trait (Vogt & Colvin, 1998) accuracy domains, the current study demonstrated no association between judges' accuracy in describing the target's personality characteristics and their accuracy in judging that person's thoughts and feelings. This novel finding suggests a limit or boundary condition to the scope of judgement ability. Although there is evidence for the good judge within both the empathic and trait accuracy domains, the good judge across these two domains does not appear to exist. An important implication is that social intelligence may be a more domain-specific ability than has hitherto often been suggested.

The Judge-Target Relationship and Judgement Accuracy

The Acquaintanceship Effect

It is important to interpret my results in light of the extensive debate over the influence of acquaintanceship on judgement validity. A major contribution of this thesis was to extend previous research by demonstrating that judgement accuracy does not plateau at the level of friendship, but instead continues to improve in the context of intimate romantic relationships. The most parsimonious explanation for this finding is that dating partners, compared to friends and strangers, possess more elaborate and rich pre-existing knowledge structures or theories concerning the target's personality and their relationship (via extensive observation and feedback). Love does not seem to be quite as blind as Shakespeare, along with many psychologists, have claimed.

I examined the links between acquaintance and accuracy cross-sectionally. However, it is important to consider how the link develops over the relationship life cycle in both the empathic and trait accuracy domains. It seems likely that there is a curvilinear association between relationship length and empathic accuracy, with empathic accuracy increasing during the process of acquaintanceship (Stinson & Ickes, 1992; Study 1 of this thesis), peaking during the early years of marriage (Bissonnette, Rusbult, & Kilpatrick, 1997), and then declining during the mature stage of the marital life cycle (Thomas et al., 1997). Thomas et al. (1997) found evidence that the lower empathic accuracy obtained by older married couples (married for 20-40 years) was produced by their lower levels of

motivation to resolve their disputes and their tendency to assume that they knew in advance what their partner was thinking and feeling.

In contrast, it is more likely that the link between trait accuracy and relationship length resembles the shape of the typical learning curve, with trait accuracy increasing markedly during the early to moderate stages of acquaintanceship (e.g., Funder & Colvin, 1988), increasing at a lesser rate between friendship and romantic relationships (Study 2 of this thesis), and perhaps peaking during the early years of marriage. Given the inherent stability and resilience of personality traits across time, it is unlikely that trait accuracy will decline in the context of well-seasoned relationships. However, such conjecture awaits further research.

The Conceptualisation of Acquaintanceship: Relationship Length vs. Relationship Closeness

An important issue that has been neglected in the accuracy literature concerns the conceptualisation of acquaintanceship. This is surprising given that research findings vary depending on the way in which acquaintanceship has been defined. The strongest evidence for the acquaintanceship effect has typically been found in studies that compared judges' accuracy across different types of judge-target relationship (e.g., friends vs. strangers, Colvin & Funder, 1991; Funder & Colvin, 1988). In such cases, acquaintanceship can be conceptualised in terms of differences in either relationship longevity or relationship closeness. However, a number of recent studies that have examined the effect of relationship length on judgement accuracy within particular types of judge-target relationships (e.g., friendship or marriage) have found equivocal results at best (Bernieri et al., 1994; Park et al., 1997; Swann & Gill, 1997).

The current study was able to disentangle relationship longevity from relationship closeness by comparing the performance of dating partners with that of friends. Dating partners had known the target for a shorter period of time than friends, but reported having a closer and more intimate relationship with the target. Given that dating partners outperformed friends, the interpretation of this result would obviously be different depending on how acquaintanceship was conceptualised. Relationship researchers have typically regarded longevity as a poor barometer of relationship development. A good number of long-term relationships (e.g., schoolmates, neighbours, and cousins) become

fixated at low levels of intimacy, or at low levels of interdependence (Berscheid, et al., 1989; Levinger, 1980). In my view acquaintanceship is more appropriately defined in terms of closeness rather than in terms of longevity.

The Interaction Between Relationship-Based Moderators and Both Person-Based and Trait-Based Moderators of Judgement Accuracy

A striking feature of my results was the pivotal role played by the unique relationship between the judge and the target in determining judgement performance in both the empathic and trait accuracy arenas. Moreover, I have argued that such an effect should most clearly emerge in the context of behaviourally rich dyadic interactions that provide well-acquainted judges with the opportunity to effectively employ their extensive pre-existing theories of the target and relationship. It is illuminating to contrast these results with those of previous research that has confined its attention to empathic judgements made by strangers, typically in the context of highly scripted and mundane interactions. The patterns of findings differ markedly in three major ways.

First, prior research in the empathic accuracy domain has found compelling evidence for the existence of the good target when judged by relative strangers (e.g., Marangoni et al., 1995). However, in the current study I found no evidence for a target effect comparing the inferences made across dating partners, friends, and relative strangers. This different result can be elegantly explained in terms of the differential reliance on data and theory in making judgements at different levels of acquaintanceship (Thomas & Fletcher, 1997). Strangers are likely to rely heavily on observational data, whereas well-acquainted judges seem to be guided by a combination of their pre-existing knowledge structures and the behavioural data evinced during the interaction.

Second, although evidence for the good judge has been found in the context of rich and complex interactions, regardless of the level of acquaintanceship, the predictors of judgement ability in this study were influenced to an important extent by the nature of the judge-target relationship. For example, my results showed that the major determinants of empathic accuracy attained by well-acquainted judges were the relationship-level variables, such as shared cognitive focus, relationship closeness, and prior problem-related disclosure. Moreover, the effects of the individual difference

variables of attributional complexity and verbal intelligence were moderated by the amount of prior knowledge or information available to the judge concerning the target.

Finally, in the trait accuracy domain, both this study and previous research (e.g., Paunonen, 1989) have found that trait observability interacts with the level of acquaintanceship between the judge and the target. Strangers can only accurately and consensually judge observable traits (e.g., traits relating to intellect in this thesis). However, well-acquainted judges are capable of accurately attributing both behavioural traits that are overtly expressed and more cognitively-based and covert personality traits.

One upshot of this discussion is that there are dangers in generalising results from research that use designs in which strangers make judgements during highly scripted interactions. The inclusion of the moderating effect of the judge-target relationship into the empirical and theoretical mix produces a revealing and more complex set of outcomes. Admittedly, given the design of this thesis it is difficult to cleanly separate the influence of the judge-target relationship from the difficulty and the richness of the judgement task. Hence, one area for future research will be to disentangle the possible effect of these two moderators on empathic accuracy (i.e., task difficulty and relationship between judge and target).

Theory-Driven Accuracy

A dominant theme of this thesis is the differential reliance on theory versus data in making judgements by insiders and outsiders. Well-acquainted judges appear to process target-related information as subjective but highly knowledgeable insiders – not as dispassionate outsiders. However, available theories in the accuracy arena as yet do not adequately deal with the complexities of such relationship-level social cognitive processes.

The role of theory-driven processing in social perception has been variously construed in the accuracy literature. Cronbach and his colleagues (Cronbach, 1955; Gage & Cronbach, 1955) viewed accuracy driven by behavioural sources as substantive and theoretically interesting, whereas theory-driven accuracy was relegated to the lowly status of pre-existing stereotypes and biases that confound the measurement of pure accuracy. Similarly, advocates of the ecological approach to social perception assert that, for the most part, targets' behaviour provides sufficient information to yield accurate

judgements, without the need for extensive cognitive processing (e.g., McArthur & Baron, 1983).

Recent theories dealing with judgement accuracy, such as RAM (Funder, 1995) and WAM (Kenny, 1994) are largely silent on the role and functions of pre-existing knowledge structures or theories. Given the belligerent undertones of the acronyms - RAM and WAM - it is perhaps ironic that both models adopt a relatively passive view of the social perceiver.¹⁸ In both theories perceivers could be construed as “arm-chair” judges, achieving greater accuracy through passive observation of an increasing number of behavioural cues (Colvin et al., 1997). In both models, the primary focus is on the nature and quantity of the behavioural data evinced by the target, not the complex theory-driven processes that occur within the head of the judge. The judges’ task is viewed as detecting and correctly utilising the incoming behavioural information, not as actively integrating such data with elaborate pre-existent theories that are both general and target/relationship-specific.

In contrast to RAM and WAM, social cognition theories assume that people cannot attend to or process the welter of data in the social environment, and therefore develop abstract knowledge structures, schemas, or lay theories to cope with this barrage of information. Once established, these rich and complex cognitive structures play a massive role in the detection and interpretation of incoming data. Indeed, the research literature dealing with close relationships overwhelmingly demonstrates the powerful role exerted by dispositional cognitive structures on relationship judgements (see, for example, Fletcher & Thomas, 1996). However, it is not plausible (and nor do I suggest) that well-acquainted judges disregard behavioural data (indeed as evidenced by the results in this thesis). Instead, prior specific theories will direct attention to certain kinds of behavioural information, guide the way this information is processed, and influence the manner in which it is used to generate social judgements. Moreover, evidence from this thesis suggests that such theories are effective devices in producing increasingly accurate readings of other people’s minds.

There are a several plausible reasons why empathic and personality judgements that are increasingly theory-guided should generally be more accurate. First, such theories are

¹⁸ During discussion with David Funder, he informed me that the acronym RAM was eminently preferable to RAT – Realistic Accuracy Theory.

typically based on extensive observation and interaction with the target both across time, and across a wide range of situations. Second, a key source of lay theories, especially in intimate relationships, consists of the provision of diagnostic feedback and disclosure concerning the target's innermost thoughts, feelings, wishes, and so forth. Finally, the psychological (and even physical) costs and benefits associated with the validity of relationship judgements are obviously much higher in close, as opposed to distant, relationships. Accordingly, individuals should be highly motivated in close relationships to forge and maintain target and relationship-specific theories that facilitate accurate judgements of their partners' underlying traits and mental states.

Alternatively, it can be (and often is) argued that such high stakes in close relationships motivate people to sacrifice accuracy goals in favour of maintaining a positive and optimistic view of one's partner and relationship. The centrality of the need to maintain a positive relationship account has been cleverly demonstrated in several recent studies. For example, Simpson et al. (1995) created an experimental setting in which dating couples experienced severe levels of threat to their relationship. The authors found that when participants who viewed their relationship as close but insecure and strongly situationally threatened, they seemingly tried to protect themselves by denying or otherwise failing to acknowledge their partners' true feelings of attraction to alternative partners. In short, partners with closer and more intimate relationship-specific theories were apparently motivated to be inaccurate in their empathic inferences (for similar results regarding personality judgements, see Murray & Holmes, 1996).

Arguments concerning these two motivational sets (accuracy vs. positivity-maintenance) have been widespread in social psychology (see Fletcher & Thomas, 1996; Thomas & Fletcher, 1997). In my view, the most plausible resolution to this debate is that both motivational sets are used in close relationships, sometimes at the same time, sometimes at different times. For example, under conditions of high threat to the relationship, the need to maintain positive theories about the relationship may supersede the need to seek the truth, and therefore may hinder judgement accuracy (e.g., Simpson et al., 1995). However, under less threatening conditions, partners are likely to weigh up the relevant evidence in a more even-handed fashion, and, hence, be more motivated to attain accurate relationship judgements.

However, it is important to note that the problem-solving context within which judgements were made in the current research was probably fairly high in threat. The fact that I found that increasingly theory-guided judgements were associated with greater accuracy suggests that that under most normal conditions lay relationships theories can, and often are, used to good effect.

Caveats and Conclusions

The current research has yielded a set of findings that is consistent with previous research but also breaks new ground and challenges some widely held propositions. This study replicated the general finding that increased acquaintanceship enhances both empathic and trait accuracy, as well as provided strong evidence for the existence of the good judge of strangers' thoughts and feelings during behaviourally rich and complex interactions. Perhaps the most original and important contribution of this thesis was the clear evidence for the major role played by the judge-target relationship in moderating judgement accuracy. First, both trait and empathic accuracy continued to increase in the context of intimate romantic relationships, rather than peaking at the level of platonic friendship. Second, relationship-level variables were the best predictors of empathic accuracy attained by well-acquainted judges. Third, the influence of the target on empathic accuracy was substantially moderated by the nature of the judge-target relationship. Finally, even when strong individual differences in judgement ability were demonstrated across different levels of acquaintanceship, the characteristics of the good judge (attributional complexity and IQ) were moderated by the level of acquaintanceship between the judge and the target.

Of course, the inevitable caveats should be noted. First, due to the nature of correlational research, one must be cautious in drawing strong causal inferences from these findings. Second, the use of a single criterion, such as self-other agreement to measure the accuracy of personality judgements could be considered problematic (although it should be noted I also assessed consensus judgements in this study, which are sometimes used as an alternative criterion of accuracy). From a realistic perspective, the criterion problem can only be effectively addressed by the use of other relevant criteria that achieve results which converge in meaningful ways (Funder, 1995). Hence,

an important task for future research is to examine the accuracy of trait judgements in relationship settings using behavioural prediction.

Moreover, the criterial status of self-descriptions of thoughts and feelings, although generally regarded in the accuracy literature as superior to that of self-judgements of personality, is also fallible. The process of self-attributions is commonly held in mainstream psychology to be itself a theoretically mediated process that is not error-free (Fletcher, 1995). I believe the most reasonable stance is that people's self-reports of their own thoughts and feelings provide good ballpark estimates of at least part of their conscious on-line experiences, but such introspective accounts should not necessarily be treated as the gold standard (not that such a single gold standard exists).

In conclusion, there are two commonly espoused views in Social and Personality Psychology concerning the influence of intimacy on judgement accuracy. The first is that "love is blind" and leads to decreased levels of judgement accuracy. The second is that accuracy plateaus at low levels of acquaintanceship. In both the central domains of trait and empathic accuracy, my research results powerfully suggest that both of these postulates are wrong.

REFERENCES

- A.C.E.R. (1982). *Manual for the advanced AL-AQ, BL-BQ, and PL-PQ cognitive tests*. Sydney, Australia: Australian Council for Educational Research.
- Acitelli, L. A. (1992). Gender differences in relationship awareness and marital satisfaction among young married couples. *Personality and Social Psychology Bulletin*, 18, 102-110.
- Acitelli, L. A., Douvan, E., & Veroff, J. (1993). Perceptions of conflict in the first year of marriage: How important are similarity and understanding? *Journal of Social and Personal Relationships*, 10, 5-19.
- Acitelli, L. K., & Young, A. M. (1996). Gender and thought in relationships. In G. J. O. Fletcher & J. Fitness (Eds.), *Knowledge structures in close relationships: A social psychological approach* (pp. 147-168). Hillsdale, NJ: Erlbaum.
- Aizen, I., & Fishbein, M. (1977). Attitude-behaviour relations: A theoretical analysis and review of empirical research. *Psychological Bulletin*, 84, 888-918.
- Akert, R. M., & Panter, A. T. (1988). Extraversion and the ability to decode nonverbal communication. *Personality and Individual Differences*, 9, 965-972.
- Allport, G. W. (1937). *Personality: A psychological interpretation*. New York: Henry Holt.
- Allport, G. W., & Odbert, H. S. (1936). Traitnames: A psycho-lexical study. *Psychological Monographs*, 47, 171.
- Altman, I., & Taylor, D. (1973). *Social Penetration: The development of interpersonal relationships*. New York: Holt, Reinhart, & Winston.
- Ambady, N., & Rosenthal, R. (1992). Thin slices of expressive behaviour as predictors of interpersonal consequences: A meta analysis. *Psychological Bulletin*, 111, 256-274.
- Andersen, S. M. (1984). Self-knowledge and social inference; II. The diagnosticity of cognitive/affective and behavioural data. *Journal of Personality and Social Psychology*, 46, 294-307.
- Aron, A., & Aron, E. N., & Smollan, D. (1992). Inclusion of Other in the Self Scale and the structure of interpersonal closeness. *Journal of Personality and Social Psychology*, 63, 596-612.

- Asch, S. E. (1946). Forming impressions of personality. *Journal of Abnormal and Social Psychology*, 41, 258-290.
- Bakan, D. (1966). *The duality of human existence*. Chicago, IL: Rand McNally, & Company.
- Baldwin, M. W. (1992). Relational schemas and the processing of social information. *Psychological Bulletin*, 112, 461-484.
- Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51, 1173-1182.
- Bernieri, F. J., Gillis, J. S., Davis, J. M., & Grahe, J. E. (1996). Dyad rapport and the accuracy of its judgement across situations: A lens model analysis. *Journal of Personality and Social Psychology*, 71, 110-129.
- Bernieri, F. J., Zuckerman, M., Koestner, R., & Rosenthal, R. (1994). Measuring person perception accuracy: Another look at self-other agreement. *Personality and Social Psychology Bulletin*, 20, 367-378.
- Berscheid, E., Snyder, M., & Omoto, A. M. (1989). The Relationships Closeness Inventory: Assessing closeness of interpersonal relationships. *Journal of Personality and Social Psychology*, 57, 792-807.
- Bissonnette, V. L., Rusbult, C. E., & Kilpatrick, S. D. (1997). Empathic accuracy and marital conflict resolution. In W. Ickes (Ed.), *Empathic accuracy* (pp. 251-281). New York: Guilford Press.
- Blackman, M., & Funder, D. (1995). *The effect of participation vs. observation on self/stranger agreement*. Paper presented at the meeting of the Western Psychological Association, Los Angeles, CA.
- Blackman, M., & Funder, D. C. (1998). The effect of information on consensus and accuracy in personality judgement. *Journal of Experimental Social Psychology*, 34, 164-181.
- Block, J., Buss, D. M., Block, J. H., & Gjerde, P. F. (1981). The cognitive style of breadth of categorization: Longitudinal consistency of personality correlates. *Journal of Personality and Social Psychology*, 40, 770-779.
- Borkenau, P., & Liebler, A. (1992). Trait inferences: sources of validity at zero acquaintance. *Journal of Personality and Social Psychology*, 62, 645-657.

- Borkenau, P., & Liebler, A. (1993a). Consensus and self-other agreement for trait inferences from minimal information. *Journal of Personality*, 61, 477-496.
- Borkenau, P., & Liebler, A. (1993b). Convergence of stranger ratings of personality and intelligence with self-ratings and measured intelligence. *Journal of Personality and Social Psychology*, 65, 546-553.
- Bruner, J. S., & Tagiuri, R. (1954). The perception of people. In G. Lindzey (Ed.), *Handbook of social psychology* (Vol. 2, pp. 634-654). Cambridge, MA: Addison-Wesley.
- Cantor, N., & Kihlstrom, J. F. (1987). *Personality and social intelligence*. Englewood Cliffs, NJ: Prentice-Hall.
- Christensen-Szalanski, J. J. J., & Beach, L. R. (1984). The citation bias: Fad and fashion in the judgement and decision literature. *American Psychologist*, 39, 75-78.
- Clark, M. S., Pataki, S. P., & Carver, V. H. (1996). Some thoughts on self-presentation of emotions in relationships. In G. J. O. Fletcher & J. Fitness (Eds.), *Knowledge structures in close relationships: A social psychological approach*. (pp. 247-274). Hillsdale, N.J.: Erlbaum.
- Cline, V. B. (1964). Interpersonal perception. In B. A. Maher (Ed.), *Progress in experimental personality research* (Vol. 1, pp. 221-284). New York: Academic Press.
- Cloyd, L. (1977). Effect of acquaintanceship on accuracy of person perception. *Perceptual and Motor Skills*, 44, 819-826.
- Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioural sciences*. 2nd Ed. New York: Erlbaum.
- Colvin, C. R. (1993a). Childhood antecedents of young adult judgeability. *Journal of Personality*, 61, 611-635.
- Colvin, C. R. (1993b). "Judgeable people": Personality, behaviour, and competing explanations. *Journal of Personality and Social Psychology*, 64, 861-873.
- Colvin, C. R., & Funder, D. C. (1991). Predicting personality and behaviour: A boundary on the acquaintanceship effect. *Journal of Personality and Social Psychology*, 60, 884-894.

- Colvin, C. R., Vogt, D., & Ickes, W. (1997). Why do friends understand each other better than strangers do? In W. Ickes (Ed.), *Empathic accuracy* (pp. 169-193). New York: Guilford Press.
- Cook, M. (1979). *Perceiving others: The psychology of interpersonal perception*. London: Methuen.
- Cronbach, L. J. (1955). Processes affecting scores on "understanding of others" and "assumed similarity". *Psychological Bulletin*, 52, 177-193.
- Cross, S. E., & Madson, L. (1997). Models of the self: Self-construals and gender. *Psychological Bulletin*, 122, 5-37.
- Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multi-dimensional approach. *Journal of Personality and Social Psychology*, 44, 113-126.
- Davis, M. H. (1994). *Empathy: A social psychological approach*. Madison, WI: Brown & Benchmark.
- Davis, M. H., & Kraus, L. A. (1997). Personality and empathic accuracy. In W. Ickes (Ed.), *Empathic accuracy* (pp. 144-168). New York: Guilford Press.
- Davis, M. H., & Oathout, A. H. (1985). Maintenance of satisfaction in romantic relationships: Empathy and relational competence. *Journal of Personality and Social Psychology*, 53, 307-410.
- Dindia, K., & Allen, M. (1992). Sex differences in self-disclosure: A meta-analysis. *Psychological Bulletin*, 112, 106-124.
- Eisenberg, N., & Lennon, R. (1983). Sex differences in empathy and related capacities. *Psychological Bulletin*, 94, 100-131.
- Epstein, S., & O'Brien, E. J. (1985). The person-situation debate in historical and current perspective. *Psychological Bulletin*, 98, 513-537.
- Fletcher, G. J. O. (1995). *The scientific credibility of folk psychology*. Hillsdale, NJ: Erlbaum.
- Fletcher, G. J. O., Danilovics, P., Fernandez, G., Peterson, D., & Reeder, G. D. (1986). Attributional complexity: An individual differences measure. *Journal of Personality and Social Psychology*, 51, 875-884.

- Fletcher, G. J. O., & Fitness, J. (1990). Occurrent social cognition in close relationship interaction: The role of proximal and distal variables. *Journal of Personality and Social Psychology*, 59, 464-474.
- Fletcher, G. J. O., Fitness, J., & Blampied, N. M. (1990). The link between attributions and happiness in close relationships: The role of depression and explanatory style. *Journal of Social and Clinical Psychology*, 9, 243-255.
- Fletcher, G. J. O., Grigg, F., & Bull, V. (1988). Organisation and accuracy of personality impressions: Neophytes versus experts in trait attribution. *New Zealand Journal of Psychology*, 17, 68-77.
- Fletcher, G. J. O., & Kininmonth, L. (1991). Interactions in close relationships and social cognition. In G. J. O. Fletcher & F. D. Fincham (Eds.), *Cognition in close relationships* (pp. 235-255). Hillsdale, NJ: Erlbaum.
- Fletcher, G. J. O., Reeder, G. D., & Bull, V. I. (1990). Bias and accuracy in trait attribution: The role of attributional complexity. *Journal of Experimental Social Psychology*, 26, 275-288.
- Fletcher, G. J. O., Rhodes, G., Rosanowski, & Lange, C. (1992). Accuracy and speed of causal processing: Experts versus novices in social judgement. *Journal of Experimental Social Psychology*, 28, 320-338.
- Fletcher, G. J. O., & Thomas, G. (1996). Close relationships lay theories: Their structure and function. In G. J. O. Fletcher & J. Fitness (Eds.), *Knowledge structures in close relationships: A social psychological approach*. (pp. 3-24). Hillsdale, N.J.: Erlbaum.
- Fletcher, G. J. O., Thomas, G., & Durant, R. (in press). Cognitive and behavioural accommodation. *Journal of Social and Personal Relationships*.
- Funder, D. C. (1983). The "consistency" controversy and the accuracy of personality judgements. *Journal of Personality*, 51, 346-359.
- Funder, D. C. (1995). On the accuracy of personality judgement: A realistic approach. *Psychological Review*, 102, 652-670.
- Funder, D. C. (1999). *Personality judgement: A realistic approach to person perception*. New York: Academic Press.

- Funder, D. C., & Colvin, C. R. (1988). Friends and strangers: acquaintanceship, agreement, and the accuracy of personality judgement. *Journal of Personality and Social Psychology*, 55, 149-158.
- Funder, D. C., & Colvin, C. R. (1991). Explorations in behavioural consistency: Properties of persons, situations, and behaviours. *Journal of Personality and Social Psychology*, 60, 773-794.
- Funder, D. C., & Colvin, C. R. (1997). Congruence of self and other's judgements of personality. In R. Hogan, J. A. Johnstone, & S. R. Briggs (Eds.), *Handbook of personality psychology* (pp. 617-647). San Diego, C.A: Academic Press.
- Funder, D. C., & Dobroth, K. M. (1987). Differences between traits: Properties associated with inter-judge agreement. *Journal of Personality and Social Psychology*, 52, 409-418.
- Funder, D. C., Kolar, D. C., & Blackman, M. C. (1995). Agreement among judges of personality: Interpersonal relations, similarity, and acquaintanceship. *Journal of Personality and Social Psychology*, 69, 656-672.
- Funder, D. C., & Ozer, D. J. (1983). Behaviour as a function of the situation. *Journal of Personality and Social Psychology*, 44, 107-112.
- Funder, D. C., & West, S. G. (1993). Consensus, self-other agreement, and accuracy in personality judgement: An introduction. *Journal of Personality*, 61, 457-467.
- Gage, N. L. (1952). Judging interests from expressive behaviour. *Psychological Monographs*, 66, No. 350: 20.
- Gage, N. L., & Cronbach, L. J. (1955). Conceptual and methodological problems in interpersonal perception. *Psychological Review*, 62, 411-422.
- Gergen, K. J. (1989). Social Psychology and the wrong revolution. *European Journal of Social Psychology*, 19, 463-484.
- Gesn, P. R. (1995). *Shared knowledge between same-sex friends: Measurement and validation*. Unpublished master's thesis, University of Texas at Arlington.
- Goldberg, L. R. (1992). The development of markers for the Big-5 factor structure. *Psychological Assessment*, 4, 26-42.
- Gottman, J. M. (1979). *Marital interaction: Experimental investigations*. New York: Academic Press.

- Graham, T. (1994). Gender, relationship, and target differences in empathic accuracy. Unpublished master's thesis: University of Texas at Arlington.
- Graham, T., & Ickes, W. (1997). When women's intuition isn't greater than men's. In W. Ickes (Ed.), *Empathic accuracy* (pp. 117-144). New York: Guilford Press.
- Hancock, M., & Ickes, W. (1996). Empathic accuracy: When does the perceiver-target relationship make a difference? *Journal of Social and Personal Relationships*, 13, 179-199.
- Hastie, R., & Rasinski, K. A. (1988). The concept of accuracy in social judgement. In K. Bar-Tal & A. W. Kruglanski (Eds.), *The social psychology of knowledge* (pp. 193-208). Cambridge, England: Cambridge University Press.
- Hendrick, C., & Hendrick, S. S. (1988). Lovers wear rose coloured glasses. *Journal of Social and Personal Relationships*, 5, 161-184.
- Hodges, S. D., & Wegner, D. M. (1997). Automatic and controlled empathy. In W. Ickes (Ed.), *Empathic accuracy* (pp. 194-218). New York: Guilford Press.
- Hogan, R. (1969). Development of an empathy scale. *Journal of Consulting and Clinical Psychology*, 33, 307-316.
- Ickes, W. (1993). Empathic accuracy. *Journal of Personality*, 61, 587-610.
- Ickes, W., Stinson, L., Bissonnette, V., & Garcia, S. (1990). Naturalistic social cognition: Empathic accuracy in mixed-sex dyads. *Journal of Personality and Social Psychology*, 59, 730-742.
- Jackson, D. N. (1982). Some preconditions for valid person perception. In M. P. Zanna, E. T. Higgins, & C. P. Herman (Eds.), *Consistency in social behaviour: The Ontario Symposium* (Vol. 2). Hillsdale, NJ: Erlbaum.
- John, O. P., & Robins, R. W. (1993). Determinants of interjudge agreement on personality traits: The Big-Five domains, observability, evaluativeness, and the unique perspective of the self. *Journal of Personality*, 61, 521-551.
- John, O. P., & Robins, R. W. (1994). Accuracy and bias in self perception: Individual differences in self-enhancement and the role of narcissism. *Journal of Personality and Social Psychology*, 66, 206-219.
- Jones, E. E. (1985). Major developments in social psychology during the past five decades. In G. Lindzey & E. Aronson (Eds.), *The handbook of social psychology* (3rd edition) (Vol. 1, pp. 47-107). New York: Random House.

- Kelly, H. H., Berscheid, E., Christensen, A., Harvey, J., Huston, T. L., Levinger, G., McClintock, E., Peplau, L. A., & Peterson, D. R. (1983). Analysing close relationships. In H. H. Kelly, A. E. Berscheid, A. Christensen, J. Harvey, T. L. Huston, G. Levinger, E. McClintock, L. A. Peplau, & D. R. Petersen (Eds.), *Close relationships* (pp. 20-67). San Francisco: Freeman.
- Kelly, A. E., & McKillop, K. J. (1996). Consequences of revealing personal secrets. *Psychological Bulletin*, 120, 450-465.
- Kenny, D. A. (1991). A general model of consensus and accuracy in interpersonal perception. *Psychological Review*, 98, 155-163.
- Kenny, D. A. (1993). A coming of age for research on interpersonal perception. *Journal of Personality*, 61, 789-807.
- Kenny, D. A. (1994). *Interpersonal perception*. New York: Guilford Press.
- Kenny, D. A., & Albright, L. (1987). Accuracy in interpersonal perception: A social relations analysis. *Psychological Bulletin*, 102, 390-402.
- Kenny, D. A., & Kashy, D. A. (1994). Enhanced co-orientation in the perception of friends: A social relations analysis. *Journal of Personality and Social Psychology*.
- Kenny, D. L., Albright, T. E., Malloy, T. E., & Kashy, D. A. (1994). Consensus in interpersonal perception: Acquaintance and the Big-Five. *Psychological Bulletin*, 116, 245-258.
- Kenrick, D. T., & Funder, D. C. (1988). Profiting from controversy: Lessons from the person-situation debate. *American Psychologist*, 43, 23-34.
- Kenrick, D. T., & Stringfield, D. O. (1980). Personality traits and the eye of the beholder: Crossing some traditional philosophical boundaries in the search for consistency in all of the people. *Psychological Review*, 87, 88-104.
- Kruglanski, A. W. (1989). The psychology of being "right": The problem of accuracy in social perception and cognition. *Psychological Bulletin*, 106, 395-409.
- Lennon, R., & Eisenberg, N. (1987). Gender and age differences in empathy and sympathy. In N. Eisenberg & J. Strayer (Eds.), *Empathy and its development* (pp. 195-217). New York: Cambridge University Press.
- Levenson, R. W., & Ruef, A. M. (1992). Empathy: A physiological substrate. *Journal of Personality and Social Psychology*, 63, 234-246.

- Levinger, G. (1980). Towards the analysis of close relationships. *Journal of Experimental Social Psychology*, 16, 510-544.
- Lopes, L. A. (1991). The rhetoric of irrationality. *Theory and Psychology*, 1, 65-82.
- Marangoni, C., Garcia, S., Ickes, W., & Teng, G. (1995). Empathic accuracy in a clinically relevant setting. *Journal of Personality and Social Psychology*, 68, 854-869.
- Markus, H., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, 98, 224-253.
- McAdams, D. P. (1988). Personal needs and personal relationships. In S. Duck, & D. F. Hay (Eds.) *Handbook of personal relationships: Theory, research, and interventions*. Chichester, UK: Wiley.
- McArthur, L. Z., & Baron, R. M. (1983). Toward an ecological theory of social perception. *Psychological Review*, 90, 215-238.
- McCrae, R. R. (1982). Consensual validation of personality traits: Evidence from self-reports and ratings. *Journal of Personality and Social Psychology*, 43, 293-303.
- Mortimer, D. C. (1996). "Reading" ourselves and "reading" others: Actual versus self-estimated empathic accuracy. Unpublished master's thesis, University of Texas at Arlington.
- Moskowitz, D. S., & Schwarz, J. C. (1982). Validity comparison of behaviour counts and ratings by knowledgeable informants. *Journal of Personality and Social Psychology*, 42, 518-528.
- Mischel, W. (1968). *Personality and assessment*. New York: Wiley.
- Murray, S. L., & Holmes, J. G. (1993). Seeing virtues in faults: Negativity and the transformation of interpersonal narratives in close relationships. *Journal of Personality and Social Psychology*, 65, 702-722.
- Murray, S. L., & Holmes, J. G. (1996). The construction of relationship realities. In G. J. O. Fletcher & J. Fitness (Eds.), *Knowledge structures in close relationships: A social psychological approach* (pp. 147-168). Hillsdale, NJ: Erlbaum.
- Noller, P. (1980). Misunderstandings in marital communication: A study of couples' nonverbal communication. *Journal of Personality and Social Psychology*, 39, 1135-1148.

- Noller, P., & Ruzzene, M. (1991). Communication in marriage: Influence of affect and cognition. In G. J. O. Fletcher & F. D. Fincham (Eds.), *Cognition in close relationships* (pp. 203-233). Hillsdale, NJ: Erlbaum.
- Noller, P., & Venardos, C. (1986). Communication awareness in married couples. *Journal of Social and Personal Relationships*, 3, 31-42.
- Norman, W. T. (1963). Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. *Journal of Abnormal and Social Psychology*, 66, 574-583.
- Park, B., Kraus, S., & Ryan, C. (1997). Longitudinal changes in consensus as a function of acquaintance and agreement in liking. *Journal of Personality and Social Psychology*, 72, 604-616.
- Paulhus, D. L., & Bruce, M. N. (1992). The effect of acquaintanceship on the validity of personality impressions: A longitudinal study. *Journal of Personality and Social Psychology*, 63, 816-824.
- Paulhus, D. L., & Reynolds, S. (1995). Enhancing target variance in personality impressions: Highlighting the person in person perception. *Journal of Personality and Social Psychology*, 69, 1233-1242.
- Paunonen, S. V. (1989). Consensus in personality judgements: Moderating effects of target-rater acquaintanceship and behaviour observability. *Journal of Personality and Social Psychology*, 56, 823-833.
- Pennebaker, J. W., Barger, S. D., & Tiebout, J. (1989). Disclosure of traumas and health among Holocaust survivors. *Psychosomatic Medicine*, 51, 577-589.
- Regan, A. M., & Hill, C. E. (1992). Investigation of what clients and counselors do and don't say in brief therapy. *Journal of Counseling Psychology*, 39, 168-174.
- Reis, H., & Patrick, B. C. (1996). Attachment and intimacy: Component processes. In E. T. Higgins, & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles*. (pp. 523-563). New York: Guilford Press
- Reis, H. T., & Shaver, P. (1988). Intimacy as an interpersonal process. In S. Duck, D. F. Hay, S. E. Hobfoll, W. Ickes, & B. M. Montgomery (Eds.) *Handbook of personal relationships* (pp. 367-389). New York: Wiley.

- Rusbult, C. E., Verette, J., Whitney, G. A., Slovik, L. F., & Lipkus, I. (1991). Accommodation processes in close relationships: Theory and preliminary evidence. *Journal of Personality and Social Psychology*, 60, 53-78.
- Schneider, D. J., Hastorf, A. H., & Ellsworth, P. C. (1979). *Person perception* (2nd ed.). Reading, MA: Addison-Wesley.
- Sillars, A. L., Pike, G. R., Jones, T. S., & Murphy, M. A. (1984). Communication and understanding in marriage. *Human Communication Research*, 10, 317-350.
- Sillars, A. L., & Scott, M. D. (1983). Interpersonal perception between intimates: An integrative review. *Human Communication Research*, 10, 153-176.
- Simpson, J. A., Ickes, W., & Blackstone, T. (1995). When the head protects the heart: Empathic accuracy in dating relationships. *Journal of Personality and Social Psychology*, 69, 629-641.
- Smither, S. (1977). A reconsideration of the developmental study of empathy. *Human Development*, 20, 253-276.
- Stelmachers, Z. T., & McHughes, R. B. (1964). Contribution of stereotyped and individualized information to predictive accuracy. *Journal of Consulting Psychology*, 28, 234-242.
- Sternberg, R. J., & Barnes, M. L. (1985). Real and ideal others in romantic relationships: Is four a crowd? *Journal of Personality and Social Psychology*, 49, 1586-1608.
- Sternberg, R. J., & Smith, C. (1985). Social intelligence and decoding skills in nonverbal communication. *Social Cognition*, 3, 168-192.
- Stinson, L., & Ickes, W. (1992). Empathic accuracy in the interactions of male friends versus male strangers. *Journal of Personality and Social Psychology*, 62, 787-797.
- Surra, C. A., & Ridley, C. A. (1991). Multiple perspectives on interaction: Participants, peers and observers. In B. M. Montgomery & S. Duck (Eds.), *Studying interpersonal interaction* (pp. 35-55). New York: Guilford Press.
- Swann, W., & Gill, M. J. (1997). Confidence and accuracy in person perception: Do we know what we think we know about our relationship partners? *Journal of Personality and Social Psychology*, 73, 747-757.
- Taft, R. (1955). The ability to judge people. *Psychological Bulletin*, 52, 1-23.

- Tetlock, P. E., & Boettger, R. (1989). Accountability: A social magnifier of the dilution effect. *Journal of Personality and Social Psychology*, 57, 388-398.
- Thomas, G., & Fletcher, G. J. O. (1997). Empathic accuracy in close relationships. In W. Ickes (Ed.), *Empathic accuracy* (pp. 194-218). New York: Guilford Press.
- Thomas, G., Fletcher, G. J. O., & Lange, C. (1997). On-line empathic accuracy in marital interaction. *Journal of Personality and Social Psychology*, 72, 839-850.
- Trope, Y., & Higgins, E. T. (1993). The what, when, and how of dispositional inference: New answers and new questions. *Personality and Social Psychology Bulletin*, 19, 493-500.
- Vangelisti, A. L. (1994). Family secrets: Forms, functions and correlates. *Journal of Social and Personal Relationships*, 11, 113-135.
- Vernon, P. E. (1933). Some characteristics of the good judge of personality. *Journal of Social Psychology*, 4, 42-57.
- Vogt, D. S., & Colvin, C. R. (1998). *The good judge of personality: Gender differences, personality correlates, and Cronbachian "artifacts"*. Unpublished Manuscript, Northeastern University.
- Watson, D., & Clark, L. A. (1991). Self versus peer ratings of specific emotional traits: Evidence of convergent and discriminant validity. *Journal of Personality and Social Psychology*, 60, 927-940.
- Wegner, D. M., & Erber, R. (1992). The hyperaccessibility of suppressed thoughts. *Journal of Personality and Social Psychology*, 63, 903-912.
- Wilson, T. D., & Schooler, J. W. (1991). Thinking too much: Introspection can reduce the quality of preferences and decisions. *Journal of Personality and Social Psychology*, 60, 181-192.

Appendix A

Big-Five Varimax-Rotated Factor Loadings for 20 Bipolar Scales

Scale	I	II	III	IV	V
Factor I. Extraversion					
Introv.-extrav.	.86*	.01	.02	.15	-.02
Unassert.-assert.	.54*	.19	-.09	.50	.09
Silent-talkat.	.83*	.01	.01	.15	.01
Unenthus.-enthus.	.37	.21	.48*	.20	.27
Factor II. Agreeableness					
Unkind-kind	.21	.70*	.16	.07	.15
Uncoop.-coop.	-.02	.72*	.15	.15	-.03
Rude-polite	-.09	.72*	.22	.05	.10
Stingy-generous	.53*	.50	-.17	.01	-.03
Factor III. Conscientiousness					
Disorg.-organ.	.01	.14	.79*	.15	.02
Irresp.-resp.	.02	.34	.76*	.10	.06
Careless-thorough	.18	.16	.72*	.06	.16
Frivolous-serious	-.04	.06	.56*	-.26	.21
Factor IV. Emotional Stability					
Tense-relaxed	.14	-.17	.18	.65*	-.10
Insecure-secure	.11	.14	.07	.75*	-.03
Emot.-unemot.	-.53*	-.17	-.38	.34	-.04
Guiltridd.-guiltfree	.05	.06	.11	.68*	.18
Factor V. Intellect					
Unintell.-intell.	.02	.28	-.09	.23	.72*
Unreflect.- reflect.	.02	.26	.11	-.24	.70*
Unimag.-imag.	.02	-.38	.30	.03	.60*
Unsophist.-sophist.	.33	.29	.17	.16	.37*

Note. Values equal to or larger than .30 are listed in boldface type.

* Highest factor loading of each scale.